

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
3. Juli 2003 (03.07.2003)

PCT

(10) Internationale Veröffentlichungsnummer  
**WO 03/054189 A2**

- (51) Internationale Patentklassifikation<sup>7</sup>: C12N 15/00 (81) Bestimmungsstaaten (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (21) Internationales Aktenzeichen: PCT/EP02/14302
- (22) Internationales Anmelddatum:  
16. Dezember 2002 (16.12.2002)
- (25) Einreichungssprache:  
Deutsch
- (26) Veröffentlichungssprache:  
Deutsch
- (30) Angaben zur Priorität:  
101 63 161.8 20. Dezember 2001 (20.12.2001) DE
- (71) Anmelder (*für alle Bestimmungsstaaten mit Ausnahme von US*): SUNGENE GMBH & CO. KGAA [DE/DE]; Corrensstrasse 3, 06466 Gatersleben (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (*nur für US*): BIESGEN, Christian [DE/DE]; Reichenstr. 29, 06484 Quedlinburg (DE).
- (74) Anwalt: BIEBERBACH, Andreas; BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).

- (84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.



A2

WO 03/054189

(54) Title: METHODS FOR THE TRANSFORMATION OF VEGETAL PLASTIDS

(54) Bezeichnung: VERFAHREN ZUR TRANSFORMATION VON PFLANZLICHEN PLASTIDEN

(57) Abstract: The invention relates to novel methods for the production of transgenic plants with genetically modified plastids, in addition to transgenic plants produced according to said method.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft neue Verfahren zur Herstellung transgener Pflanzen mit gentechnisch veränderten Plastiden sowie die mit diesen Verfahren hergestellten transgenen Pflanzen.

BEST AVAILABLE COPY

**Verfahren zur Transformation von pflanzlichen Plastiden****Beschreibung****5**

Die vorliegende Erfindung betrifft neue Verfahren zur Herstellung transgener Pflanzen mit gentechnisch veränderten Plastiden sowie die mit diesen Verfahren hergestellten transgenen Pflanzen.

**10** Ziel biotechnologischer Arbeiten an Pflanzen ist die Herstellung von Pflanzen mit vorteilhaften, neuen Eigenschaften zum Beispiel zur Steigerung der landwirtschaftlichen Produktivität, zur Qualitätssteigerung bei Nahrungsmitteln oder zur Produktion bestimmter Chemikalien oder Pharmazeutika.

**15**

Plastiden stellen Organellen innerhalb pflanzlicher Zellen mit einem eigenen Genom dar. Sie haben eine wesentliche Funktion in der Photosynthese sowie der Aminosäure- und Lipidbiosynthese. Das Genom der Plastiden besteht aus einer doppelsträngigen,

**20** zirkulären DNA mit einer durchschnittlichen Größe von 120 bis 160 kb und liegt - zum Beispiel in Blattzellen - mit ca. 1900 bis 50000 Kopien pro Zelle vor (Palmer (1985) Ann Rev Genet 19:325-54). Ein einzelnes Plastid trägt dabei eine Kopienzahl von ca. 50 bis 100. Der Begriff der Plastiden umfasst Chloroplasten,

**25** Proplastiden, Etioplasten, Chromoplasten, Amyloplasten, Leukoplasten und Elaioplasten (Heifetz P (2000) Biochimie 82:655-666). Die verschiedenen Formen sind in einander überführbar und entstehen alle aus den Proplastiden. Daher enthalten alle Ausprägungsformen der Plastiden die gleiche Erbinformation. Bevorzugt werden in der Literatur jedoch grüne Zellen, die Chloroplasten als Ausprägungsform enthalten, als Ausgangsmaterial für die Transformation von Plastiden genutzt.

Es ist für die Pflanzenbiotechnologie von großem wirtschaftlichen Interesse, effiziente Methoden für die Plastidentransformation zu entwickeln (McFadden G (2001) Plant Physiol. 125:50-53). Die stabile Transformation von Plastiden höherer Pflanzen gehört dabei zu den großen Herausforderungen.

**40** Die bei der Insertion in die nukleäre DNA oft angewandte Technik der ungezielten (illegitimen) DNA-Insertion haben bei der Plastidentransformation den Nachteil, dass mit sehr hoher Wahrscheinlichkeit auf dem gedichten plastidären Genom ein essentielles Gen getroffen wird, was häufig letal für die Pflanze wäre. In Plastiden ist daher eine gezielte Insertion von Fremd-DNA vorteilhaft.

Verschiedene Verfahren zur gezielten Insertion in das plastidäre Genom sind beschrieben. Zuerst wurde die Plastidentransformation in Grünalgen beschrieben (Boynton JE et al. (1988) Science 240: 1534-1538; Blowers AD et al. (1989) Plant Cell 1:123-132), später 5 in höheren Pflanzen wie Tabak (Svab Z et al. (1990) Proc Natl Acad Sci USA 87:8526-8530).

EP-A 0 251 654, US 5,932,479, US 5,451,513, US 5,877,402,  
WO 01/64024, WO 00/20611, WO 01/42441, WO 99/10513, WO 97/32977,  
10 WO 00/28014, WO 00/39313 beschreiben Methoden und DNA-Konstrukte  
zur Transformation von Plastiden höherer Pflanzen, wobei die  
zu transformierende DNA über homologe Rekombination ("Doppel-  
Crossover") in das Plastom (Plastidengenom) eingebaut wird. Im  
allgemeinen werden auf jeder Seite der zu insertierenden Sequenz  
15 homologe Regionen von 1000 bp oder mehr genutzt. Dies führt sehr  
schnell zu großen Vektoren, deren Handhabung wenig komfortabel  
ist. Darüber hinaus sinkt die Effizienz der Transformation.  
Mit zunehmender Länge der zu integrierenden Fremd-DNA sinkt die  
Effizienz der homologen Rekombination. Ein weiterer Nachteil  
20 ist die Tatsache, dass für jede Pflanzenart ein homologer Bereich  
identifiziert werden muss, der für den Prozess der DNA Inte-  
gration mittels Doppel-Crossover genutzt werden kann. WO 99/10513  
beansprucht, eine intergenische DNA Sequenz identifiziert zu  
haben, die ausreichend homolog zwischen den Genomen der Chlоро-  
25 plasten vieler höherer Pflanzen sei und so als universelle  
Zielsequenz fungieren kann. Dass dieser Vektor bei anderen  
Arten als Tabak erfolgreich genutzt werden kann, wurde indies  
nicht nachgewiesen. Im Gegenteil: In WO 01/64024 passt derselbe  
Erfinder den Transformationsvektor an Nicht-Tabak Pflanzenarten  
30 an, indem er aus diesen isolierte homologe DNA-Sequenzen ver-  
wendet. Da bei allen der oben beschriebenen Verfahren nur wenige  
Rekombinationsergebnisse resultieren, ist eine Selektion der  
rekombinanten plastidären DNA-Moleküle erforderlich.

35 Die Plastiden-DNA höherer Pflanzen liegt in bis zu mehreren  
tausend Kopien pro Zelle vor. Um eine stabile Integration der  
Fremd-DNA zu erreichen, müssen alle Kopien der Plastiden-DNA  
gleichermaßen verändert werden. Im Fall der Plastidentrans-  
formation spricht man davon, den homotransplastomen Zustand  
40 erreicht zu haben. Diesen Zustand erreicht man durch einen  
sogenannten "segregation and sorting" Prozess, indem man die  
Pflanzen unter ständigen Selektionsdruck hält. Durch den an-  
haltenden Selektionsdruck werden während der Zell- und Plastiden-  
teilung solche Plastiden angereichert, in denen bereits viele  
45 Kopien der plastidären DNA verändert wurden. Der Selektions-  
druck wird aufrecht gehalten, bis der homotransplastome Zu-  
stand erreicht ist (Guda C et al. (2000) Plant Cell Reports

19:257-262). Es ist eine große Herausforderung, alle Kopien des Plastidengenoms zu verändern, um homotransplastome Pflanzen zu erhalten, die auch über Generationen ohne Zugabe eines Selektionsmittels das Fremdgen stabil in ihr plastidäres Genom eingebaut haben (Bogorad L (2000) TIBTECH 18:257-263). Zusätzlich zu dem ständigen Selektionsdruck wird gegebenenfalls durch wiederholte Regeneration bereits transformierten Gewebes das Erreichen des homotransplastomen Zustand sichergestellt (Svab Z und Maliga P (1993) Proc Natl Acad Sci USA 90:913-917). Dies 10 jedoch schränkt das Pflanzenmaterial, welches für die Plastidentransformation zur Verfügung steht, ein. Es wird daher vorgeschlagen, das Transgen ggf. mit einem anderen Gen, welches essentiell für das Überleben der Pflanze ist, zu koppeln. Gewebekultur-Techniken und Selektionsprozesse sind meist nicht 15 universell auf alle Pflanzenarten anwendbar und stellen eine wesentliche Limitierung der Plastidentransformation dar, die insbesondere die Übertragbarkeit der Methode auf andere Arten als Tabak betrifft (Kota M et al. (1999) Proc Natl Acad Sci USA 96:1840-1845). Eine kürzlich veröffentlichte Transformation von Plastiden von Tomate basiert auf Modifikationen 20 im Regenerations- und Selektionsschema (Ruf S et al. (2001) Nature Biotech 19:870-875), die jedoch kosten- und zeitintensiv sind. Ein anderer Ansatz zielt darauf ab, die Anzahl der Plastiden pro Zelle und der DNA Moleküle pro Plastid zu 25 reduzieren, um so weniger DNA Moleküle modifizieren zu müssen (Bogorad L (2000) TIBTECH 18:257-263). Insgesamt sind die Selektions- und Segregationsprozesse sehr zeitaufwendig.

WO 99/10513 beschreibt ein Verfahren bei dem auf dem zu transformierenden Plasmid ein plastidärer ORI (origin of replication, Replikationsursprung) lokalisiert wird, um so die Anzahl der für die Integration in das Plastidengenom zur Verfügung stehenden Kopien des zu transformierenden Vektors zu erhöhen (Guda C et al. (2000) Plant Cell Reports 19:257-262).

35 Die Notwendigkeit die Technik der Plastidentransformation zu verbessern wird auch bei Heifetz und Tuttle erwähnt (Heifetz P und Tuttle AM (2001) Curr Opin Plant Biol 4:157-161). WO 00/32799 lehrt, die Effizienz der Plastidentransformation durch den Einsatz von Pflanzen mit vergrößerten Plastiden zu erhöhen. Dadurch ergibt sich eine große Oberfläche der Plastiden, durch die die zu transformierende DNA leichter in die Plastiden eindringen kann. Der Mechanismus der DNA Integration erfolgt aber auch hier mittels konventioneller homologer Rekombination wie bei den oben 40 beschriebenen Verfahren.

Verschiedene weitere Verfahren zur sequenzspezifischen Integration von DNA - insbesondere in die nukleäre DNA - sind beschrieben. Beschrieben ist ein Verfahren basierend auf selbst-spleißenden Introns der Gruppe II. Selbstspeißende Introns der 5 Gruppe II können sequenzspezifisch - beispielsweise in intronlose Gene - insertieren. Die sequenzspezifische Hydrolyse der Ziel-DNA wird von einem RNA-Protein (Ribonukleoprotein) Komplex katalysiert. Die Sequenzspezifität der Endonuklease-Funktion wird dabei vor allem durch Ausbilden von Basenpaarungen zwischen dem RNA- 10 Anteil des Ribonukleoprotein-Komplexes und der Ziel-DNA bestimmt. Die Verwendung von Introns der Gruppe II als Vektoren für Fremd-DNA wurde diskutiert. Durch Modifikation bestimmter Sequenzen eines Gruppe II Introns konnte dessen Ziel-Spezifität verändert werden. Auch konnten weitere Sequenzen in Gruppe II Introns 15 insertiert werden, ohne deren Funktionen zu zerstören (Yang J. et al. (1996) Nature 381:332-335; Eickbush TH (1999) Curr Biol 9:R11-R14; Matsuura M et al. (1997) Genes Develop 11:2910-2924; Cousineau B et al. (1998) Cell 94: 451-462). Die Anpassung an bestimmte Zielsequenzen und die Ermittlung der damit verbundenen 20 Regeln ist jedoch aufwendig und wurde bislang im Detail nur für das L1.ltrB Intron entschlüsselt (Mohr G et al. (2000) Genes Develop 14:559-573). Zudem war die Effizienz des Retrohomings durch die Modifikation signifikant reduziert und nicht jedes der getesteten modifizierten Introns insertierte in die gewünschte 25 Ziel-DNA. Nachteil der Technik ist, dass einige Positionen in der Nukleotid-Sequenz festgelegt sind, was die Auswahl der Ziel-Region in der zu transformierenden DNA einschränkt (Guo H et al. (1997) EMBO J 16:6835-6848). Außerdem scheint die Effizienz des Retrohoming-Prozesses gegenüber dem der Wildtyp-Introns ver- 30 mindert. Die Effizienz der Introninsertion an verschiedenen Stellen an den betrachteten Genen war verschieden hoch. Ziel der Arbeiten war die Bereitstellung einer verbesserten Methode zur gezielten Insertion von DNA in die nukleäre DNA von Organismen, die keine effiziente homologe Rekombination erlauben (Guo et al. 35 (2000) Science 289:452-456). Die beschriebenen Versuche sind extrachromosomal sowohl in dem Prokaryonten *E.coli* als auch in humanen Zellen realisiert worden. Die Übertragbarkeit auf die chromosomale DNA höherer Organismen oder die Anwendbarkeit auf plastidäre DNA wurde weder beschrieben noch gezeigt. Es wurde 40 lediglich vorgeschlagen, zu versuchen, dieses System derart zu optimieren, dass eine Insertion in chromosomal DNA höherer Eukaryoten erfolgen kann. Dieses System soll eine alternative Methode in höheren Eukaryoten darstellen, die keine effiziente homologe Rekombination zeigen (Guo et al. (2000) Science 45 289:452-456). Dies gilt nicht für Plastiden höherer Pflanzen, bei denen die homologe Rekombination - zumindest bei einzelnen

plastidären DNA-Molekülen - in der Regel unproblematisch realisiert werden kann.

Neben Tabak wurde die Plastidentransformation in Kartoffel  
5 (Sidorov VA et al. (1999) Plant J 19:209-216; WO 00/28014),  
Petunie (WO 00/28014), Reis (Khan MS und Maliga P (1999) Nature  
Biotech 17:910-915; WO 00/07431; US 6,153,813), Arabidopsis  
(Sikdar SR et al. (1998) Plant Cell Reports 18: 20-24;  
WO 97/32977) und Raps (WO 00/39313) gezeigt (Übersichtsartikel:  
10 Bogorad L (2000) TIBTECH 18:257-263). Kürzlich wurden auch  
transplastome Tomatenpflanzen beschrieben (Ruf S et al. (2001)  
Nature Biotech 19:870-875).

Die Erzeugung sequenzspezifischer Doppelstrangbrüche mit Hilfe  
15 von Restriktionsenzymen in eukaryontischen Genomen u.a. Pflanzen  
ist beschrieben (Puchta H (1999) Methods Mol Biol 113:447-451).

WO 96/14408 beschreibt die Homing-Restriktionsendonuklease I-SceI  
sowie verschiedene Möglichkeiten deren Verwendung. Eine Anwendung  
20 zur Insertion von DNA-Sequenzen in plastidäre DNA ist nicht  
beschrieben.

Posfai et al. beschreiben ein Verfahren zum Austausch von Genen  
in dem Prokaryoten E.coli (Posfai G et al. (1999) Nucleic Acids  
25 Res 27(22):4409-4415). Dabei kommt es im E.coli Genom zu einer  
intramolekularen Rekombination zwischen dem endogenen und dem  
mutierten Gen, die durch einen Schnitt mit dem Restriktionsenzym  
I-SceI induziert wird. Rekombinationen in E.coli verlaufen deut-  
lich effizienter und nach vermutlich anderen Mechanismen als im  
30 Kern höherer Eukaryonten (zum Beispiel beschrieben bei Kuzminov A  
(1999) Microbiol Mol Biol Rev. 63(4):751-813).

"Homing" meint das Phänomen, dass in einem Kompartiment zwei oder  
mehr Kopien einer DNA-Sequenz existieren, wobei wenigstens eine  
35 dieser beiden Sequenzen durch eine weitere DNA-Sequenz unter-  
brochen ist, und in Folge eine Kopie der unterbrechenden DNA-  
Sequenz auch in die nicht unterbrochene DNA-Sequenz eingebracht  
wird. Meist tritt das Phänomen als Intron-Homing auf. Hier  
existieren zwei oder mehr Allele eines Gens in einem Komparti-  
40 ment, wobei wenigstens eines dieser Allele kein Intron besitzt.  
In Folge wird eine Kopie des Introns auch in das intronlose Allel  
eingebracht.

Introns in plastidären Genen höherer Pflanzen sind beschrieben  
45 (Vogel J et al. (1999) Nucl Acids Res 27:3866-3874; Jenkins BD  
et al. (1997) Plant Cell 9:283-296; Xu MQ et al. (1990) Science  
250: 1566-1570). Das Spleißen eines arteigenen, nicht modi-

fizierten Introns mit den natürlichen Exonbereichen an einem ektopischen Lokus im Plastidengenom ist ebenfalls beschrieben (Bock R und Maliga P (1995) Nucl Acids Res 23(13):2544-2547). Es gibt bisher keine Versuche, artfremde Introns, die darüber hinaus derart modifiziert sind, dass sie zusätzliche genetische Information besitzen und/oder in einer nicht natürlichen Sequenzumgebung spleißen, in Plastiden höherer Pflanzen einzubringen.

Versuche von Eddy und Gold zum "Homing"-Prozess in *E.coli* zeigten, das bestimmte Rekombinationssysteme erforderlich sind. Die Art des Rekombinationssystems des Wirtes ist eine Schlüsselvariable (Eddy SR und Gold L (1992) Proc Natl Acad Sci USA 89:1544-1547). Es konnte daher nicht davon ausgegangen werden, dass der natürlich vorkommende Prozess des "Homings" beliebig von einem in einen anderen Organismus übertragen werden kann, zumal wenn dieser Prozess in diesem Organismus natürlicherweise mutmaßlich nicht vorkommt.

Dürrenberger et al. beschreiben die Induktion einer intrachromosomal Rekombination in Chloroplasten der einzelligen Grünalge *Chlamydomonas reinhardtii* unter Verwendung der I-CreI Homing-Endonuklease (Dürrenberger F et al. (1996) Nucleic Acid Res 24(17):3323-3331). Die Rekombination erfolgt zwischen dem endogenen 23S-Gen und einer ins Chromosom eines I-CreI Deletionsstammes insertierten 23S-cDNA, die eine I-CreI Schnittstelle enthält. Doppelstrangbrüche werden durch "Mating" des entsprechenden transgenen Organismus mit einem I-CreI natürlicherweise exprimierenden Organismus induziert. Zum Zeitpunkt des Doppelstrangbruches ist die Fremd-DNA bereits in die chromosomale DNA insertiert und die Rekombination erfolgt intramolekular und nicht zwischen zwei separaten Molekülen.

Kürzlich wurde gezeigt, dass ein in *Chlamydononas reinhardtii* natürlich vorkommendes, mobiles Intron, welches auch eine Homing-Endonuklease kodiert, effizient in eine intronlose Kopie transformiert werden kann (Odom OW et al. (2001) Mol Cell Biol 21: 3472-3481). Die Steigerung der Transformationsrate war dabei von der Anwesenheit der Homing-Endonuklease abhängig. In der Diskussion wird allgemein und ohne konkrete Ausführungshinweise vorgeschlagen, die Plastidentransformation dadurch zu verbessern, dass Doppelstrangbrüche induziert werden. Dazu sollen zunächst die Erkennungsregionen selten schneidender Nukleasen in einem ersten Schritt eingebracht werden, und das nachfolgende Integrationsereignis dann an dem selben Locus stattfinden. Genaue Hinweise, wie die Erkennungsregionen einzubringen sind, welche Nukleasen und Erkennungsregionen verwendet werden können, wie der erste und der zweite Schritt konkret gestaltet werden etc. werden

nicht gegeben. Bisher wurde lediglich gezeigt, dass in Plastiden der Alge *Chlamydomonas* das Einbringen eines arteigenen Introns mittels der natürlicherweise mit der Mobilität des Intron verbundenen Homing Endonuklease funktionierte. Zudem wurden die Ergebnisse in einer Algenart generiert. Die oben erwähnten Versuche von Eddy und Gold mit *E. coli*, bei denen wie in Plastiden höherer Pflanzen keine mobilen Introns der Gruppe I bekannt sind, zeigen, dass eine Übertragbarkeit auf heterologe Systeme nicht ohne weiteres möglich ist. Die Übertragbarkeit der Beobachtungen von der Alge *Chlamydomonas* auf höhere Pflanzen ist für den Fachmann daher keinesfalls naheliegend. Im Gegenteil gibt es zahlreiche Hinweise, die die Übertragbarkeit eher in Frage stellen:

1. Homing-Systeme sind nicht einfach von einem auf das andere System übertragbar (Eddy SR und Gold L (1992) Proc Natl Acad Sci USA 89:1544-1547). Die Übertragbarkeit auf höhere Pflanzen ist um so mehr fraglich, da in den bereits sequenzierten Plastidengenomen höherer Pflanzen ([http://megasun.bch.umontreal.ca/ogmp/projects/other/cp\\_list.html](http://megasun.bch.umontreal.ca/ogmp/projects/other/cp_list.html)) 20 keine Homing Endonukleasen indentifiziert wurden. Es ist daher davon auszugehen, dass die im Plastidengenom höherer Pflanzen gefundenen Introns nicht mobil sind und dort natürlicherweise kein Homing-Mechanismus besteht.
2. *Chlamydomonas* hat nur jeweils ein Plastid, während in den Zellen höherer Pflanzen bis zu 100 Plastiden pro Zelle vorliegen. 25
3. Die konventionelle Plastidentransformation ist in *Chlamydomonas* um Größenordnungen effizienter als in höheren Pflanzen, was darauf hindeutet, dass diese beiden Systeme nicht direkt miteinander vergleichbar sind. Hinsichtlich der Regeneration von transplastomen Algen bzw. transplastomen Pflanzen könnte auch eine entscheidende Rolle spielen, dass die Teilung der Plastiden der Algen mit dem Zellzyklus synchronisiert ist, 30 während dies nicht für die Plastiden der höheren Pflanzen gilt (Sato N (2001) Trends Plant Science 6:151-155).
4. Die Mechanismen der DNA Integration in Plastiden von *Chlamydomonas* und höheren Pflanzen scheinen grundlegend verschieden zu sein. So wurde gefunden, dass die zwischenartliche Plastidentransformation (nutzen von homologen Bereichen anstelle von identischen Sequenzen) in *Chlamydomonas* zu einer deutlichen Reduktion der Transformations- 45 effizienz führt, was in Tabak aber nicht beobachtet werden kann. Ähnliches gilt auch für den Abstand eines molekularen Markers auf der homologen DNA zur heterologen Sequenz auf dem

- Transformationsplasmid: Je näher der molekulare Marker am Rand der Zielregion für die Integration mittels Doppel-Crossover liegt, desto seltener wird er bei der Transformation in Plastiden von Chlamydomonas übertragen. In Tabak wurden multiple Rekombinationsmechanismen beobachtet, wobei aber auch molekulare Marker nahe am Rand der homologen Bereiche effektiv bei der Transformation in das Plastidengenom übertragen wurden (Kavanagh TA et al. (1999) Genetics 152: 1111-1122 und Referenzen darin).
5. In *Chlamydomonas* verschmelzen die Plastiden der beiden Eltern während einer Kreuzung, auch bei einer zwischenartlichen Kreuzung. Die Plastidenfusion ist bei *Chlamydomonas* ein natürlicher Prozess und auch die DNA der Plastiden wird gemischt und neu rekombiniert. Daher machen mobile Introns in den Organellen dieser Organismen Sinn. Im Gegensatz dazu werden die Plastiden bei den meisten höheren Pflanzen uniparental vererbt, so dass es weder zu einem Mischen der plastidären DNA kommt, noch Rekombinationen zwischen der plastidären DNA der Mutter und des Vaters auftreten können. Selbst in den Pflanzenarten, in denen es zu biparentaler Vererbung der Plastiden kommt, konnte keine Fusion der Plastiden beobachtet werden. Es liegt daher die Vermutung nahe, dass eine natürliche Plastidenfusion in höheren Pflanzen ausgeschlossen ist (Hagemann R (1992) Plastid genetics in higher plants; in Cell Organelles, Herausgeber: Herrmann RG, Springer Verlag, Wien, S.65-96) und das Mechanismen wie das Intron-Homing nicht entwickelt sind oder sogar unterdrückt werden.
30. Die Erhöhung der Effizienz der homologen Rekombination innerhalb der Kern-DNA mit Hilfe von selten schneidenden Endonukleasen wurde bei verschiedenen Organismen beschrieben (Puchta H et al. (1993) Nucleic Acids Research. 21(22):5034-40; Puchta H et al. (1996) Proc Natl Acad Sci USA 93:5055-5060; Rong YS und Golic KG (2000) Science 28:2013-2018; Jasin M (1996) Trends Genet 12: 224-229). Im Unterschied zu Plastiden ist die Insertion mittels homologer Rekombination in die Kern-DNA problematisch und erfolgt meist mittels zufälliger illegitimen Integration. Dies zeigt, dass im Kern-Genom etablierte Techniken nicht unbedingt auf die Plastiden übertragen werden können. Bei Plastiden höherer Pflanzen findet im Gegensatz zu der Situation im Kern nahezu ausschließlich und mit hoher Effizienz eine Integration über homologe Rekombination statt (Bock R und Hagemann R (2000) Progress in Botany 61:76-90; Maliga P et al. (1994) Homologous recombination and integration of foreign DNA in plastids of higher

plants. In Homologous recombination and gene silencing in plants.  
Paszkowski J, ed. (Kluwer Academic publishers), S.83-93).

Die Effizienz der homologen Rekombination zur DNA-Integration  
5 in das Plastom wurde allgemein nicht als limitierender Faktor  
angesehen und - im Gegenteil - als unkritisch eingestuft. Die  
aktuelle Forschung zur Optimierung der Plastidentransformation  
richtet sich demzufolge nicht auf eine Optimierung der homologen  
Rekombination, sondern zum Beispiel auf verbesserte Selektions-  
10 marker, verbesserte Selektions- und Regenerationstechniken etc.  
Nichtsdestotrotz sind bislang keine wesentlichen Durchbrüche  
erzielt worden.

Wie die oben beschriebenen Verfahren und Probleme bei der  
15 Plastidentransformation deutlich hervorheben, ist die Bereit-  
stellung neuer Verfahren zur Herstellung homotransplastomer  
Pflanzen ein lange bestehendes ungelöstes Bedürfnis der Pflanzen-  
biotechnologie. Ein weiteres Bedürfnis ist die Vermeidung von  
Antibiotika- oder Herbizidselektionsmarkern aus Gründen der  
20 Zulassung und Verbraucherakzeptanz. Für die Plastidentrans-  
formation ist bisher noch kein Verfahren beschrieben, dass ohne  
einen solchen Selektionsmarker auskommt.

Es ergab sich daher die Aufgabe, neue Verfahren zu entwickeln,  
25 die eine effiziente Integration von Fremd-DNA in alle Kopien der  
Plastiden-DNA sicher stellt und eine effiziente Selektion ent-  
sprechender homotransplastomer Pflanzen ermöglicht. Diese Aufgabe  
wurde durch Bereitstellung des erfindungsgemäßen Integrations-/  
Selektionsverfahrens in überraschender Weise gelöst.

30 Ein erster Gegenstand der Erfindung betrifft ein Verfahren  
zur Integration einer DNA-Sequenz in die plastidäre DNA einer  
mehrzelligen Pflanze oder von dieser abgeleiteten Zelle und zur  
Selektion überwiegend homotransplastomer Zellen oder Pflanzen,  
35 dadurch gekennzeichnet, dass

- a) die plastidären DNA-Moleküle besagter mehrzelliger  
Pflanze oder von dieser abgeleiteten Zelle mindestens  
eine Erkennungssequenz zur gezielten Induktion von DNA-  
40 Doppelstrangbrüchen enthalten und
- b) mindestens ein Enzym geeignet zur Induktion von DNA-  
Doppelstrangbrüchen an der Erkennungssequenz zur gezielten  
Induktion von DNA-Doppelstrangbrüchen sowie mindestens ein  
45 Transformationskonstrukt umfassend eine Insertionssequenz  
in mindestens einem Plastid besagter mehrzelliger Pflanze

10

oder von dieser abgeleiteten Zelle zusammengebracht werden,  
und

- 5 c) die Induktion von DNA-Doppelstrangbrüchen an den Erkennungssequenzen zur gezielten Induktion von DNA-Doppelstrangbrüchen erfolgt, und
- 10 d) die Insertionssequenz in die plastidäre DNA insertiert, wobei die Funktionalität der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen deaktiviert wird, so dass besagte Erkennungssequenz nicht mehr durch das Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen geschnitten werden kann, und
- 15 e) Pflanzen oder Zellen isoliert werden, bei denen die Insertionssequenz in die plastidären DNA-Moleküle insertiert wurde.

Das System ermöglicht überraschenderweise eine erhebliche  
20 Steigerung der Effizienz bei der Herstellung überwiegend homotransplastomer Pflanzen. Dabei wird sowohl die Effizienz der Insertion in die plastidäre DNA als auch die Effizienz des Selektionsprozess überwiegend homotransplastomer Pflanzen gesteigert.

25 Durch die Anwendung des erfindungsgemäßen Verfahrens ergibt sich ein Selektionsdruck, die Insertionssequenz in alle Kopien der plastidären DNA einzubauen. Im Idealfall erfolgt die Verbreitung der Insertionssequenz unabhängig von Selektionsmarkern wie Herbizid- oder Antibiotikaresistenzen. Dies hat deutliche Vorteile bezüglich Zulassung- und/oder Verbraucherakzeptanz. Die Verwendung solcher Selektionsmarker kann jedoch die Effizienz weiter steigern. Das erfindungsgemäße Verfahren zur Herstellung homoplastomer Pflanzen ist den im Stand der Technik beschriebenen 30 deutlich überlegen, da es einen schnelleren, effizienteren und damit kostengünstigeren Weg zu homotransplastomen Pflanzen bietet. Ferner hat das System den Vorteil, dass die zur Transformation genutzten Konstrukte klein gehalten werden können, da im Vergleich zur Integration mittels des Doppel-Crossovers die 35 homologen Bereiche im Plastidentransformations-Vektor kleiner sein bzw. gänzlich fehlen können.

## 11.

Gegenüber der Transformation des Zellkerns hat die Transformation von Plastiden zahlreiche Vorteile. Unter anderem sind zu nennen:

- a) Während die homologe Rekombination in die nukleäre DNA nur schwer zu realisieren ist, kann in Plastiden DNA leicht an einem vordefinierten Ort mittels Doppel-Crossover, einer Form der homologen Rekombination, integriert werden. Positions-effekte oder "Genesilencing", die man bei Kerntransformationen aufgrund der illegitimen Integration an einen nicht vordefinierten Lokus findet, werden so vermieden.
- b) Es können sehr hohe Expressionslevel erreicht werden, vermutlich aufgrund der hohen Kopiezahl der plastidären DNA.
- c) Die DNA der Plastiden wird bei höheren Pflanzen in der Regel nur maternal vererbt, so dass die eingebrachte Fremd-DNA nicht über Pollen verbreitet und ein Auskreuzen somit effektiv unterbunden werden kann.
- d) Die prokaryotische Natur der Plastiden ermöglicht die Expression von Genen im Rahmen einer polycistronischen Operonstruktur. Daher ist es nicht notwendig, jedes zu exprimierende Gen mit einem eigenen Promotor etc. auszurüsten. Dies erleichtert das Einbringen vieler Gene auf einmal, etwa um ganze Biosynthesewege in die Plastiden einzubringen.

"Plastid" meint die Proplastiden sowie alle daraus hervorgehenden Organellen wie beispielsweise Chloroplasten, Etioplasten, Chromoplasten, Amyloplasten, Leukoplasten, Dermaplasten und Elaioplasten (Heifetz P (2000) Biochimie 82:655-666).

"Plastom" meint das Genom, also die Gesamtheit der genetischen Information, eines Plastids.

"Homotransplastom" meint einen transplastomen und homoplastomen Zustand.

"Transplastom" meint in Bezug auf beispielsweise eine Pflanze, Zelle, Gewebe, Plastid oder plastidäre-DNA alle solche durch gentechnische Methoden zustande gekommene Formen der vorgenannten, die eine plastidäre DNA umfassen, die durch gentechnische Methoden modifiziert wurde, wobei die Modifikation beispielhaft Substitutionen, Additionen, Deletionen, Inversion oder Insertionen eines oder mehrerer Nukleotidreste umfassen kann.

## 12

"Heteroplastom" meint das Vorliegen einer gemischten Population verschiedener plastidärer DNAs innerhalb eines einzelnen Plastides oder innerhalb einer Population von Plastiden innerhalb einer pflanzlichen Zelle oder Gewebe.

5

"Homoplastom" meint eine einheitliche Population von plastidärer DNA innerhalb eines einzelnen Plastides oder innerhalb einer Population von Plastiden innerhalb einer pflanzlichen Zelle oder Gewebe. Homoplastome Zellen, Gewebe oder Pflanzen sind genetisch

10 stabil, da sie nur eine Art plastidärer DNA umfassen d.h. sie bleiben im allgemeinen homoplastom auch wenn der Selektionsdruck nicht mehr anhält. Durch Selbstung erhaltene Nachkommen sind ebenfalls homoplastom.

15 Im Rahmen dieser Erfindung meint "überwiegend homoplastom" oder "überwiegend homotransplastom" all solche Pflanzen oder Zellen, bei denen der Anteil der erwünschten plastidären hinsichtlich eines Merkmals veränderten DNA-Moleküle - beispielsweise mit der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrang-

20 brüchen oder der insertierten Insertionssequenz - mindestens 50 %, bevorzugt mindestens 70 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt mindestens 95 % von der Gesamtheit aller plastidären DNA-Moleküle in einer Pflanze oder einem Gewebe, Zelle bzw. Plastid derselben beträgt. Überwiegend

25 homoplastome oder überwiegend homotransplastome Pflanzen können durch weiteres Aufrechterhalten des Selektionsdruckes und gegebenenfalls wiederholenden Regenerationen in homoplastome oder homotransplastome Pflanzen umgewandelt werden. Aufgrund des "Homing"-Prozesses ist jedoch ein andauernder Selektionsdruck

30 nicht zwingend erforderlich. In einer besonderen Ausführungsform ist daher eine überwiegend homoplastome bzw. homotransplastome Pflanze rein homoplastom bzw. homotransplastom. Eine Pflanze, die bezüglich einer DSB-Erkennungssequenz überwiegend homoplastom bzw. homotransplastom oder rein homoplastom bzw. homotransplastom

35 ist, wird infolge als "Masterpflanze" bezeichnet. Der Anteil der erwünschten plastidären hinsichtlich eines Merkmals veränderten DNA-Moleküle kann beispielsweise mittels Southern Analyse - wie beispielhaft in Beispiel 4 beschrieben - in der dem Fachmann bekannten Weise ermittelt werden. Das Verhältnis zwischen den plas-

40 stidären Ausgangs-DNA Molekülen und den hinsichtlich eines Merkmals veränderten plastidären DNA-Molekülen lässt sich durch Vergleich der Intensität der jeweiligen Banden bestimmen.

"Mehrzellige Pflanze oder von dieser abgeleiteten Zelle" meint 45 allgemein jede Zellen, Gewebe, Teile oder Vermehrungsgut (wie Samen oder Früchte) einer Pflanze, die in ihrem adulten Zustand einen mehrzelligen Organismus darstellt oder darstellen kann.

Eingeschlossen sind im Rahmen der Erfindung alle Gattungen und Arten höherer und niedrigerer Pflanzen des Pflanzenreiches. Einjährige, mehrjährige, monocotyledone und dicotyledone Pflanzen sind bevorzugt. Eingeschlossen sind reife Pflanzen, Saatgut,

5 Sprosse und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut (zum Beispiel Knollen, Samen oder Früchte) und Kulturen, zum Beispiel Zell- oder Kalluskulturen. Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem

10 frühen Entwicklungsstadium.

Bevorzugt sind Pflanzen nachfolgender Pflanzenfamilien: Amarantaceae, Asteraceae, Brassicaceae, Carophyllaceae, Chenopodiaceae, Compositae, Cruciferae, Cucurbitaceae, Labiateae, Leguminosae,

15 Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Rubiaceae, Saxifragaceae, Scrophulariaceae, Solanaceae, Sterculiaceae, Tetragoniacea, Theaceae, Umbelliferae.

Bevorzugte monokotyle Pflanzen sind insbesondere ausgewählt aus den monokotylen Kulturpflanzen, wie zum Beispiel der Familie der Gramineae wie Reis, Mais, Weizen oder andere Getreidearten wie Gerste, Hirse, Roggen, Triticale oder Hafer sowie dem Zuckerrohr sowie alle Arten von Gräsern.

25 Bevorzugte dikotyle Pflanzen sind insbesondere ausgewählt aus den dikotylen Kulturpflanzen, wie zum Beispiel

- Asteraceae wie Sonnenblume, Tagetes oder Calendula und andere mehr,

30

- Compositae, besonders die Gattung Lactuca, ganz besonders die Art sativa (Salat) und andere mehr,

35

- Cruciferae, besonders die Gattung Brassica, ganz besonders die Arten napus (Raps), campestris (Rübe), oleracea cv Tastie (Kohl), oleracea cv Snowball Y (Blumenkohl) und oleracea cv Emperor (Broccoli) und weitere Kohlarten; und der Gattung Arabidopsis, ganz besonders die Art thaliana sowie Kresse oder Canola und andere mehr,

40

- Cucurbitaceae wie Melone, Kürbis oder Zucchini und andere mehr,

- Leguminosae besonders die Gattung Glycine, ganz besonders die Art max (Sojabohne) Soja sowie Alfalfa, Erbse, Bohnengewächsen

45 oder Erdnuss und andere mehr

## 14

- Rubiaceae, bevorzugt der Unterklasse Lamiidae wie beispielsweise Coffea arabica oder Coffea liberica (Kaffestrauch) und andere mehr,
  
- 5 - Solanaceae besonders die Gattung Lycopersicon, ganz besonders die Art esculentum (Tomate) und die Gattung Solanum, ganz besonders die Art tuberosum (Kartoffel) und melongena (Aubergine), sowie Tabak oder Paprika und andere mehr,
  
- 10 - Sterculiaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Theobroma cacao (Kakastrauch) und andere mehr,
  
- Theaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Camellia sinensis oder Thea sinensis (Teestrauch) und
- 15 andere mehr,
  
- Umbelliferae, besonders die Gattung Daucus (ganz besonders die Art carota (Karotte)) und Apium (ganz besonders die Art graveolens dulce (Selarie)) und andere mehr; und die Gattung
- 20 Capsicum, ganz besonders die Art annum (Pfeffer) und andere mehr,
  
- sowie Lein, Soja, Baumwolle, Hanf, Flachs, Gurke, Spinat, Möhre, Zuckerrübe und den verschiedenen Baum-, Nuss- und Weinarten,
- 25 insbesondere Banane und Kiwi.

Umfasst sind ferner Schmuckpflanzen, Nutz- oder Zierbäume, Blumen, Schnittblumen, Sträucher oder Rasen. Beispielhaft aber nicht einschränkend seien zu nennen Angiospermen, Bryophyten

- 30 wie zum Beispiel Hepaticae (Leberblümchen) und Musci (Moose); Pteridophyten wie Farne, Schachtelhalm und Lycopoden; Gymnospermen wie Koniferen, Cycaden, Ginkgo und Gnetalen, die Familien der Rosaceae wie Rose, Ericaceae wie Rhododendrons und Azaleen, Euphorbiaceae wie Weihnachtssterne und Kroton, Caryophyllaceae
- 35 wie Nelken, Solanaceae wie Petunien, Gesneriaceae wie das Usambaraveilchen, Balsaminaceae wie das Springkraut, Orchidaceae wie Orchideen, Iridaceae wie Gladiolen, Iris, Freesie und Krokus, Compositae wie Ringelblume, Geraniaceae wie Geranien, Liliaceae wie der Drachenbaum, Moraceae wie Ficus, Araceae wie Philodendron
- 40 und andere mehr.

Am meisten bevorzugt sind Arabidopsis thaliana, Nicotiana tabacum, Tagetes und Brassica napus sowie alle Gattungen und Arten, die als Nahrungs- oder Futtermittel zum Einsatz kommen,

- 45 wie die beschriebenen Getreidearten, oder sich zur Herstellung von Ölen eignen, wie Ölsaaten, Nussarten, Soja, Sonnenblume, Kürbis und Erdnuss.

"Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen" (infolge "DSBI-Enzym" für "double strand-break inducing enzyme") meint allgemein all solche Enzyme, die in der Lage sind, 5 sequenzspezifisch Doppelstrangbrüche in doppelsträngige DNA zu erzeugen. Beispielsweise aber nicht einschränkend sind zu nennen:

1. Restriktionsendonukleasen, bevorzugt Typ II Restriktions-  
10 endonukleasen, besonders bevorzugt Homing-Endonukleasen wie weiter unten im Detail beschrieben.
2. Künstliche Nukleasen wie weiter unten im Detail beschrieben,  
15 wie beispielsweise chimere Nukleasen, mutierte Restriktions- oder Homing-Endonukleasen oder RNA-Proteinpartikel abgeleitet von mobilen Introns der Gruppe II.

Sowohl natürliche als auch künstlich hergestellte DSBI-Enzyme sind geeignet. Bevorzugt sind all solche DSBI-Enzyme, deren 20 Erkennungssequenz bekannt ist und die entweder in Form ihrer Proteine (beispielsweise durch Aufreinigung) gewonnen oder unter Verwendung ihrer Nukleinsäuresequenz exprimiert werden können.

Bevorzugt wird das DSBI-Enzym unter Kenntnis seiner spezifischen 25 Erkennungssequenz so ausgewählt, dass es zusätzlich zu der Ziel-Erkennungssequenz keine weitere funktionellen Erkennungsregionen im Plastidengenom besitzt. Ganz besonders bevorzugt sind deshalb Homing-Endonukleasen (Übersicht: Belfort M und Roberts RJ (1997) Nucleic Acids Res 25:3379-3388; Jasin M (1996) 30 Trends Genet 12:224-228; Internet: <http://rebase.neb.com/rebase/rebase.homing.html>; Roberts RJ and Macelis D (2001) Nucleic Acids Res 29: 268-269). Diese erfüllen aufgrund ihrer langen Erkennungssequenzen diese Anforderung. Aufgrund der geringen Größe des Plastoms ist es jedoch auch denkbar, dass DSBI-Enzyme 35 mit kürzeren Erkennungssequenzen (beispielsweise Restriktionsendonukleasen) mit Erfolg eingesetzt werden können.

Neben der oben beschriebenen bevorzugten Ausführungsform, bei der lediglich eine einzelne Erkennungssequenz für das DSBI-Enzym 40 in der plastidären DNA vorhanden ist, sind Fälle denkbar, bei denen auch weitere, funktionell gleiche Erkennungssequenzen vorteilhaft genutzt werden können. Dies ist insbesondere der Fall, wenn das Plastom duplizierte Gene (z.B. in Form von invertierten "Repeats") umfasst. Hier soll die Integration in alle Kopien 45 erfolgen, so dass auch ein Schnitt in allen Kopien wünschenswert ist.

## 16

Die für derartige Homing-Endonukleasen kodierenden Sequenzen können beispielsweise aus dem Chloroplastengenom von Chlamydomonas isoliert werden (Turmel M et al. (1993) J Mol Biol 232: 446-467). Sie sind klein (18 bis 26 kD), weisen in ihrem offenen Leseraster (ORF) eine "coding usage" auf, die direkt für nukleäre oder plastidäre Expression höheren Pflanzen (Monnat RJ Jr et al. (1999) Biochem Biophys Res Com 255:88-93) geeignet ist.

Weitere Homing-Endonukleasen sind unter der oben angegebenen Internet-Adresse aufgeführt, zu nennen sind beispielsweise Homing-Endonukleasen wie F-SceI, F-SceII, F-SuvI, F-TevI, F-TevII, I-AmaI, I-AnII, I-CeuI, I-CeuAIIP, I-ChuI, I-CmoI, I-CpaI, I-CpaII, I-CreI, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIVP, I-CsmI, I-CvuI, I-CvuAIP, I-DdIII, I-DirI, I-DmoI, I-HspNIP, I-LlaI, I-MsoI, I-NaaI, I-NanI, I-NcliP, I-NgrIP, I-NitI, I-NjaI, I-Nsp236IP, I-PakI, I-PboIP, I-PcuIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PpbIP, I-PpoI, I-SPBetaIP, I-ScaI, I-SceI, I-SceII, I-SceIII, I-SceIV, I-SceV, I-SceVI, I-SceVII, I-SexIP, I-SneIP, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquIP, I-Ssp6803I, I-StPhiJP, I-StPhiST3P, I-StPhiS3bP, I-TdeIP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPA1P, I-UarHGPA13P, I-VinIP, I-ZbiIP, PI-MtuI, PI-MtuHIP, PI-MtuHIIIP, PI-PfuI, PI-PfuII, PI-PkoI, PI-PkoII, PI-PspI, PI-Rma43812IP, PI-SPBetaIP, PI-SceI, PI-TfuI, PI-TfuII, PI-ThyI, PI-TliI, PI-TliII.

Bevorzugt sind dabei die Homing-Endonukleasen, deren Gensequenzen bereits bekannt sind, wie beispielsweise F-SceI, I-CeuI, I-ChuI, I-DmoI, I-CpaI, I-CpaII, I-CreI, I-CsmI, F-TevI, F-TevII, I-TevI, I-TevII, I-AnII, I-CvuI, I-LlaI, I-NanI, I-MsoI, I-NitI, I-NjaI, I-PakI, I-PorI, I-PpoI, I-ScaI, I-SceI, I-Ssp6803I, PI-PkoI, PI-PkoII, PI-PspI, PI-TfuI, PI-TliI.

Besonders bevorzugt werden Homing-Endonuklease genutzt, die so in der Natur, besonders bevorzugt natürlicherweise in Organellen vorkommen. Am meisten bevorzugt stammen die Homing-Endonukleasen aus Organismen, die bei ähnlichen Temperaturen leben wie Pflanzen. Hier sind von besonderem Interesse die aus Hefe und Chlamydomonas Spezies identifizierten Homing Endonukleasen. Natürlich ist es auch denkbar, Homing Endonukleasen, die aus extremophilen Organismen isoliert wurden, zu nutzen, solange diese in den Plastiden der zu transformierenden Pflanze aktiv sind.

17.

Ganz besonders bevorzugt sind

- I-CeuI (Cote MJ und Turmel M (1995) Curr Genet 27:177-183.; Gauthier A et al. (1991) Curr Genet 19:43-47; Marshall (1991) Gene 104:241-245; GenBank Acc.-No.: Z17234 Nukleotide 5102 bis 5758),
- I-ChuI (Cote V et al. (1993) Gene 129:69-76; GenBank Acc.-No.: L06107, Nukleotide 419 bis 1075),
- 10 - I-CmoeI (Drouin M et al. (2000) Nucl Acids Res 28: 4566-4572),
- I-CpaI aus *Chlamydomonas pallidostigmatica* (GenBank Acc.-No.: L36830, Nukleotide 357 bis 815; Turmel M et al. (1995) Nucleic Acids Res 23:2519-2525; Turmel, M et al. (1995) Mol Biol Evol 12:533-545; siehe auch SEQ ID NO: 13 bzw. 14)
- 15 - I-CpaII (Turmel M et al. (1995) Mol Biol Evol 12:533-545; GenBank Acc.-No.: L39865, Nukleotide 719 bis 1423),
- I-CreI (Wang J et al. (1997) Nucleic Acids Res 25: 3767-3776; Dürrenberger, F und Rochaix JD (1991) EMBO J 10:3495-3501; GenBank Acc.-No.: X01977, Nukleotide 571 bis 1062),
- 20 - I-CsmI (Ma DP et al. (1992) Plant Mol Biol 18:1001-1004)
- I-NanI (Elde M et al. (1999) Eur J Biochem. 259:281-288; GenBank Acc.-No.: X78280, Nukleotide 418 bis 1155),
- 25 - I-NitI (GenBank Acc.-No.: X78277, Nukleotide 426 bis 1163),
- I-NjaI (GenBank Acc.-No.: X78279, Nukleotide 416 bis 1153),
- 30 - I-PpoI kodiert auf der extrachromosomaler DNA im Kern von *Physarum polycephalum* (Muscarella DE und Vogt VM (1989) Cell 56:443-454; Lin J und Vogt VM (1998) Mol Cell Biol 18:5809-5817; GenBank Acc.-No.: M38131, Nukleotide 86 bis 577). Darüber hinaus kann auch die von einem alternativen Startkodon ausgehende längere Sequenz, kodierend für I-PpoI genutzt werden. Diese Sequenz umfaßt in der Sequenz gemäß GenBank Acc.-No M38131 die Nukleotide 20 bis 577. Bevorzugt wird die kürzere Sequenz genutzt; sie kann jedoch an jeder Stelle durch die entsprechend längere ausgetauscht werden.
- 35 - Besonders bevorzugt wird im Rahmen dieser Erfindung eine künstliche Sequenz genutzt, die für die selben Aminosäuren kodiert, wie die Sequenz der Nukleotide 86 bis 577 der Se-
- 40
- 45

18

quenz gemäß GenBank Acc.-No.: M38131 (siehe auch SEQ ID NO: 5, 11, 12, "70 oder 71),

- I-PspI (GenBank Acc.-No.: U00707, Nukleotide 1839 bis 3449),

5

- I-ScaI (Monteilhet C et al. (2000) Nucleic Acids Res 28: 1245-1251; GenBank Acc.-No.: X95974, Nukleotide 55 bis 465).

- 10 - I-SceI aus den Mitochondrien der Bäckerhefe (WO 96/14408; US 5,962,327 dort Seq ID NO: 1),

- Endo SceI (Kawasaki et al. (1991) J Biol Chem 266:5342-5347, identisch zu F-SceI; GenBank Acc.-No.: M63839, Nukleotide 159 bis 1589),

- 15 - I-SceII (Sarguiel B et al. (1990) Nucleic Acids Res 18:5659-5665),

- I-SceIII (Sarguiel B et al. (1991) Mol Gen Genet. 255:340-341),

20

- I-Ssp6803I (GenBank Acc.-No.: D64003, Nukleotide 35372 bis 35824),

- 25 - I-TevI (Chu et al. (1990) Proc Natl Acad Sci USA 87:3574-3578; Bell-Pedersen et al. (1990) Nucleic Acids Res 18:3763-3770; GenBank Acc.-No.: AF158101, Nukleotide 144431 bis 143694),

- 30 - I-TevII (Bell-Pedersen et al. (1990) Nucleic Acids Res 18:3763-3770; GenBank Acc.-No.: AF158101, Nukleotide 45612 bis 44836),

- I-TevIII (Eddy et al. (1991) Genes Dev. 5:1032-1041),

35 Ganz besonders bevorzugt sind kommerziell erhältliche Homing-Endonukleasen wie I-CeuI, I-SceI, I-PpoI, PI-PspI oder PI-SceI. Am meisten bevorzugt sind I-SceI und I-PpoI. Während das Gen codierend für I-PpoI in der natürlichen Form genutzt werden kann, besitzt das Gen kodierend für I-SceI eine Editierstelle. Da die 40 entsprechende Editierung in den Plastiden höherer Pflanzen im Unterschied zu den Mitochondrien der Hefe nicht durchgeführt wird, muss eine künstliche das I-SceI Protein codierende Sequenz zur heterologen Expression dieses Enzyms eingesetzt werden (US 5,866,361).

45

19

- Neben den aufgeführten Homing-Endonukleasen gibt es weitere Intron-kodierte Enzyme, die an homologen Stellen der Genome verwandter Organismen zu finden sind. Diese Homing-Endonukleasen besitzen in der Regel eine ähnliche Sequenzspezifität und sind daher ebenso gut als DSBI Enzym für die Einführung eines DSB in das Plastom an der DSB-Erkennungssequenz geeignet. So erkennen neben I-CreI die entsprechende Sequenz auch Sob2593c, Clu2593c, Col2593c, CiY2593c, Hla2593c, Cag2593c, I-CvuI, I-PakI, Tmu2593c, Msp2593c, I-MsoI, Sdu2593c, Mvi2593m, Nol2593m oder Aca2593m.
- 10 Neben I-CeuI erkennen entsprechende Sequenz auch I-CecI, I-CmoI, I-CelI, I-CpaIII, I-CmuI, I-CluI, I-SobI oder I-AstI. Neben I-CpaI erkennen eine entsprechende Sequenz auch Cbr1931c, Cfr1931c, Cme1931c, Cge1931c, Pcr1931c, Msp1931c, Msol1931c, PtU1931c, CvU1931m, Msp1931m, Msol1931m, Nol1931m, Aca1931m oder 15 Sne1931b. Darüber hinaus gibt es Introns, die an der Position 1951 der 23S rDNA insertiert sind (Nummerierung bezieht sich auf homologen Position in der 23S rDNA von *E. coli*). Auch diese Introns kodieren für putative Homing-Endonukleasen, die als DSBI Enzyme zum spezifischen Schneiden der plastidären DNA genutzt werden können. Zu diesen gehören beispielsweise Cbr1951c, 20 Msp1951c, Msol1951c, CvU1951m oder Aca1951m (Lucas P et al. (2001) Nucl Acids Res 29:960-969).

Am meisten bevorzugt sind die Homing-Endonukleasen gemäß den 25 Proteinsequenzen beschrieben durch SEQ ID NO: 5, 12 oder 14. Bei der Herstellung entsprechender Expressionskassetten werden demzufolge Nukleinsäuresequenzen eingesetzt, die für ein Protein gemäß SEQ ID NO: 5, 12 bzw. 14 kodieren, insbesonders bevorzugt ist dabei die Verwendung der Nukleinsäuresequenzen gemäß SEQ ID 30 NO: 11 bzw. 13 oder einer Expressionskassette gemäß SEQ ID NO: 4.

Die Enzyme können in der dem Fachmann geläufigen Art und Weise aus ihren Herkunftsorganismen aufgereinigt und/oder die für sie kodierende Nukleinsäuresequenz kloniert werden. Die Sequenzen 35 verschiedener Enzyme sind in der GenBank hinterlegt (s.o.).

Als künstliche DSBI-Enzyme seien beispielhaft chimäre Nukleasen zu nennen, die sich aus einer unspezifischen Nuklease domäne und einer sequenzspezifischen DNA-Bindungsdomäne (z.B. bestehend aus 40 Zinkfingern) zusammensetzen (Smith J et al. (2000) Nucl Acids Res 28(17):3361-3369; Bibikova M et al. (2001) Mol Cell Biol. 21:289-297). So wurde beispielsweise die katalytische Domäne der Restriktionsendonuklease FokI an Zinkfingerbinde-Domänen fusioniert wodurch die Spezifität der Endonuklease definiert 45 wurde (Chandrasegaran S & Smith J (1999) Biol Chem 380:841-848; Kim YG & Chandrasegaran S (1994) Proc Natl Acad Sci USA 91:883-887; Kim YG et al. (1996) Proc Natl Acad Sci USA

20

93:1156-1160). Auch die katalytische Domäne der Ho-Endonuklease aus Hefe konnte bereits mit der beschriebenen Technik eine vordefinierte Spezifität verliehen werden, indem diese an die Zinkfingerdomäne von Transkriptionsfaktoren fusioniert wurde  
5 ( Nahon E & Raveh D (1998) Nucl Acids Res 26:1233-1239).

Wie erwähnt, eignen sich insbesondere Zinkfingerproteine als DNA-Bindungsdomäne im Rahmen von chimären Nukleasen. Diese DNA-bindenden Zinkfingerdomänen können an jede beliebige DNA-Sequenz 10 angepasst werden. Entsprechende Verfahren zur Herstellung entsprechender Zinkfingerdomänen sind beschrieben und dem Fachmann bekannt (Beerli RR et al. (2000) Proc Natl Acad Sci USA 97(4):1495-1500; Beerli RR et al. (2000) J Biol Chem 275(42):32617-32627; Segal DJ and Barbas CF 3rd. (2000) Curr 15 Opin Chem Biol 4(1):34-39; Kang JS and Kim JS (2000) J Biol Chem 275(12):8742-8748; Beerli RR et al. (1998) Proc Natl Acad Sci USA 95(25):14628-14633; Kim JS et al. (1997) Proc Natl Acad Sci USA 94(8):3616-3620; Klug A (1999) J Mol Biol 293(2):215-218; Tsai SY et al. (1998) Adv Drug Deliv Rev 30(1-3):23-31; Mapp AK et al. 20 (2000) Proc Natl Acad Sci USA 97(8):3930-3935; Sharrocks AD et al. (1997) Int J Biochem Cell Biol 29(12):1371-1387; Zhang L et al. (2000) J Biol Chem 275(43):33850-33860). Verfahren zur Herstellung und Selektion von Zink-Finger DNA-Bindedomänen mit hoher Sequenzspezifität sind beschrieben. (WO 96/06166, 25 WO 98/53059, WO 98/53057). Durch Fusion einer so erhaltenen DNA-Bindedomäne an die katalytische Domäne einer Endonuklease (wie beispielsweise der FokI oder Ho-Endonuklease) kann man chimäre Nukleasen mit jeder beliebigen Spezifität herstellen, die als DSBI-Enzyme im Rahmen der vorliegenden Erfindung vorteilhaft 30 eingesetzt werden können.

Künstliche DSBI-Enzyme mit veränderter Sequenzspezifität können auch durch Mutagenese bereits bekannter Restriktionsendonukleasen oder Homing Endonukleasen durch den Fachmann geläufige Methoden 35 erzeugt werden. Insbesondere von Interesse ist neben der Mutagenese von Homing-Endonukleasen mit dem Ziel, eine veränderte Substratspezifität zu erzielen, auch die Mutagenese von Maturasen. Maturasen haben häufig viele Gemeinsamkeiten mit Homing-Endonukleasen und lassen sich durch wenige Mutationen ggf. in 40 Nukleasen umwandeln. Dies wurde beispielsweise für die Maturase im bi2 Intron der Bäckerhefe gezeigt. Lediglich zwei Mutationen in dem die Maturase kodierenden offenen Leseraster (ORF) reichten aus, diesem Enzym eine Homing-Endonuklease Aktivität zu verleihen (Szczepanek & Lazowska (1996) EMBO J 15:3758-3767).

## 21

Weitere künstliche Nukleasen können mit Hilfe mobiler Gruppe II-Introns und den von ihnen kodierten Proteinen oder Teile dieser Proteine erzeugt werden. Viele mobile Gruppe II Introns bilden mit den von ihnen kodierten Proteinen RNA-Protein-Partikel, die sequenzspezifisch DNA erkennen und schneiden können. Die Sequenzspezifität kann dabei durch Mutagenese von bestimmten Bereichen des Introns (siehe unten) den Bedürfnissen angepasst werden (WO 97/10362).

10 Dem Fachmann sind verschiedene Verfahren bekannt, um ein DSBI-Enzym in Plastiden einzubringen oder dort zu exprimieren. Beispielhaft jedoch nicht einschränkend sind zu nennen:

- a) Nukleäre Expression unter Verwendung plastidärer Transit-peptide

Eine Expressionskassette kodierend für ein DSBI-Enzym-Fusionsprotein kann in der dem Fachmann bekannten Weise konstruiert, in den Zellkern eingeführt und - optional - stabil in die chromosomale DNA integriert werden. Für einen Transport in die Plastiden wird das DSBI-Enzym bevorzugt in Fusion mit einer Plastidenlokalisationssequenz (PLS) exprimiert. Verfahren, an sich nicht in den Plastiden lokalisierte Proteine gezielt in die Plastiden zu transportieren, sowie verschiedene PLS-Sequenzen sind beschrieben (Klosgen RB und Weil JH (1991) Mol Genet 225(2):297-304; Van Breusegem F et al. (1998) Plant Mol Biol 38(3):491-496). Bevorzugt sind PLS, welche nach Translokation des DSBI-Enzyms in die Plastiden vom DSBI-Enzym-Teil enzymatisch abgespalten werden. Insbesondere bevorzugt ist die PLS, die von der plastidären Nicotiana tabacum Transketolase oder einem anderen Transitpeptid (z.B. dem Transitpeptid der kleinen Untereinheit der Rubisco oder der Ferredoxin NADP Oxidoreduktase als auch der Isopentenylpyrophosphat Isomerase-2) oder dessen funktionellem Äquivalent abgeleitet ist. Für die nukleäre Expression sind im Prinzip alle Promotoren geeignet, die eine Expression in Pflanzen ermöglichen. Beispiele sind weiter unten gegeben. Bevorzugt sind konstitutive Promotoren wie der 35S Promotor des CaMV oder der Nitrilase-1 Promotor des ntl Gens aus Arabidopsis (GenBank Acc.-No.: Y07648.2, Nukleotide 2456 bis 4340; Hillebrand H et al. (1998) Plant Mol Biol 36 (1):89-99; Hillebrand H et al. (1996) Gene 170(2):197-200).

22

Bevorzugte PLS Sequenzen sind:

- i) das Transitpeptid der Isopentenylisomerase (IPP) aus  
Arabidopsis (GenBank Acc.-No.: NC\_003074; Nukleotide  
5 604657 - 604486)
  - ii) Transitpeptide abgeleitet von der kleine Untereinheit  
(SSU) der Ribulosebisphosphatcarboxylase (Rubisco ssu)  
aus beispielsweise Erbse, Mais, Sonnenblume oder  
10 Arabidopsis.
    - Arabidopsis thaliana: GenBank Acc.-No.: z.B.  
AY054581, AY054552;
    - Erbse, GenBank Acc.-No.: z.B. X00806, Nukleotide  
15 1086 bis 1256; X04334, X04333 (Hand JM (1989)  
EMBO J 8(11):3195-206). Hier besonders bevorzugt:  
Expressionskassette und Transitpeptid (Erbse, rbcS3A)  
aus Vektor pJIT117 (Guerineau F (1988) Nucleic Acids  
20 Res 16(23):11380. Besonders bevorzugt ist die Peptid-  
sequenz gemäß SEQ ID NO: 35. Am meisten bevorzugt für  
die Verwendung zur Konstruktion von entsprechenden  
Expressionskonstrukten ist die Nukleinsäuresequenz  
gemäß SEQ ID NO: 34.
    - Mais, GenBank Acc.-No.: z.B. S42568, S42508
    - Sonnenblume, GenBank Acc.-No.: Y00431, Nukleotide 301  
25 bis 465.
  - iii) Transitpeptide abgeleitet von Genen der pflanzlichen  
Fettbiosynthese wie das Transitpeptid des plastidären  
"Acyl Carrier Protein" (ACP) (z.B. die Arabidopsis  
30 thaliana beta-ketoacyl-ACP synthetase 2; GenBank Acc.-  
No.: AF318307), die Stearyl-ACP-Desaturase,  $\beta$ -Ketoacyl-  
ACP Synthase oder die Acyl-ACP-Thioesterase (z.B.  
A.thaliana mRNA for acyl-(acyl carrier protein)thio-  
esterase: GenBank Acc.-No.: Z36911).
  - iv) das Transitpeptid der GBSSI ("Starch Granule Bound  
Synthase I")
  - v) das Transitpeptid der LHCP II Gene.
- 45 Besonders bevorzugt ist die PLS der plastidären Transketolase  
aus Tabak (SEQ ID NO: 36). Zur Expression entsprechender  
Fusionsproteine können im Rahmen dieser Erfindung ver-

schiedene PLS-Nukleinsäurekassetten in den drei Lese-  
rastern als KpnI/BamHI Fragment verwendet werden (der  
Translationsstart (ATG-Kodon) ist in der NcoI Schnittstelle  
5 lokalisiert) (pTP09 SEQ ID NO: 37; pTP10 SEQ ID NO: 38; pTP11  
SEQ ID NO: 39).

Ein weiteres Beispiel für eine vorteilhaft einzusetzende PLS  
ist das Transitpeptid der plastidären Isopentenylpyrophosphat  
Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* (SEQ ID NO: 40).  
10 Ganz besonders bevorzugt können die Nukleinsäuresequenzen  
kodierend für drei Kassetten (entsprechend den drei Lese-  
rastern) der PLS der Isopentenylpyrophosphat Isomerase-2  
(IPP-2) aus *Arabidopsis thaliana* verwendet werden (EcoRV/  
SalI-Kassetten mit ATG; IPP-9 SEQ ID NO: 41; IPP-10  
15 SEQ ID NO: 42; IPP-11 SEQ ID NO: 43).

Die erfindungsgemäßen Nukleinsäuren können synthetisch her-  
gestellt oder natürlich gewonnen sein oder eine Mischung aus  
synthetischen und natürlichen Nukleinsäure-Bestandteilen ent-  
20 halten, sowie aus verschiedenen heterologen Genabschnitten  
verschiedener Organismen bestehen.

Die für das Transitpeptid kodierende Sequenz kann ganz oder  
teilweise die Peptidsequenz des Ursprungsproteins umfassen.  
25 Eine exakte Bestimmung der für den Transport essentiellen  
Aminosäurereste ist nicht erforderlich, solange die Funktio-  
nalität der PLS – nämlich der Transport in die Plastiden –  
gewährleistet ist, und die Funktion des DSBI-Enzyms nicht  
gänzlich zerstört wird. Ganz besonders bevorzugt sind die  
30 nachfolgenden PLS-Sequenzen:

PLS1: N-MASSSSLTLSQAILSRSVPRHGSASSSQLSPSSLTFSGLKSNPNITTSSRRR  
TPSSAAAAAVVRSPAIRASAATETIEKTETAGS-C (SEQ ID NO: 36).  
Entspricht der PLS der plastidären Transketolase aus  
35 Tabak.

PLS2: N-MSASSLFNLPLIRLRLSALSSSFSSFRFAHRPLSSISPRKLPNFRAFGTAA  
MTDTKDGSRVDM-C (SEQ ID NO: 40). Entspricht der PLS  
der Isopentenylpyrophosphat Isomerase-2 (IPP-2),  
40 wobei das letzte Methionin bevorzugt das Start-  
Methionin des DSBI Enzymes ist.

Die Homing-Endonuklease gemäß SEQ ID NO: 69 stellt ein bevor-  
zugtes Fusionsproteine aus der nativen I-Ppo I Nuklease und  
45 der der IPP-Plastidenlokalisationssequenz dar. Bevorzugt

wird dieses Protein durch die Sequenz mit der SEQ ID NO: 68 kodiert.

5 Fusionsproteine aus PLS und DSBI-Enzym sind im Rahmen dieser Erfindung unter dem Begriff des DSBI-Enzyms subsumiert. Wird eine DSBI-Enzym nukleär exprimiert, so wird unter dem DSBI-Enzym bevorzugt ein Fusionsprotein aus PLS und DSBI-Enzym verstanden.

10 Ein weiterer Gegenstand der erfindung betrifft Fusionproteine aus DSBI-Enzymen mit Plastidenlokalisationssequenz (PLS) Sequenzen sowie Expressionskassetten enthaltend ein Fusionsprotein aus einer Plastidenlokalisationssequenz (PLS) und einem DSBI-Enzym unter Kontrolle eines im pflanzlichen Zellkern funktionellen Promotors. Entsprechend geeignete 15 Promotoren sind dem Fachmann bekannt und weiter unten beschrieben. Die Expressionskassette kann weitere Elemente, wie beispielsweise Transkriptionsterminatoren und/oder Selektionsmarker beinhalten.

20 b) Expression in Plastiden

Eine Expression in Plastiden kann auch durch direkte Einführung einer Expressionskassette für das DSBI-Enzym in Plastiden, ggf. Integration in die plastidäre DNA und Expression des DSBI-Enzyms erfolgen. In einer bevorzugten Ausführungsform ist diese Expressionskassette in dem die Insertionssequenz umfassenden Transformationskonstrukt enthalten.

30 Dabei können zum einen spezifische Plastiden- oder Chromoplasten-Promotoren - wie unten im Detail beschrieben - zum Einsatz kommen. Eine gezielte plastidäre Expression kann auch erreicht werden, wenn man zum Beispiel einen viralen, bakteriellen oder Bakteriophagen Promotor verwendet, die resultierende Expressionskassette in die plastidäre DNA einbringt und die Expression dann durch die korrespondierende virale, bakterielle oder Bakteriophagen RNA-Polymerase induziert. Die korrespondierende RNA-Polymerase kann wiederum auf verschiedene Art und Weise - bevorzugt durch nukleäre Transformation als Fusionsprotein mit einer PLS - in die Plastiden eingebracht werden. Entsprechende Verfahren sind beschrieben (WO 95/16783, WO 97/06250, US 5,925,806). Das Einbringen in Plastiden wird bevorzugt mittels Mikroinjektion und besonders bevorzugt mittels Partikelbeschuss durchgeführt.

25

## c) Einbringen als RNA

- Das DSBI-Enzym kann auch durch Einführung der für das DSBI-Enzym kodierenden - beispielsweise in vitro erzeugten - mRNA z.B. über Mikroinjektion, Partikel-Bombardierung (biolistische Verfahren), Polyethylenglykol- oder Liposomen-vermittelte Transfektion in Plastiden eingebracht werden. Diese Ausführungsform ist vorteilhaft, da hier keine DSBI-Enzym kodierenden Sequenzen im Plastom oder Genom verbleiben.
- Bevorzugt wird die das DSBI-Enzym kodierende RNA durch in vitro Transkription in der dem Fachmann bekannten Weise hergestellt.

## d) Einbringen als Protein

- Das DSBI-Enzym kann direkt beispielsweise über Mikroinjektion, Partikel-Bombardierung (biolistische Verfahren), Polyethylenglykol-Transfektion oder Liposomen-vermittelte Transfektion in Plastiden eingebracht werden. Diese Ausführungsform ist vorteilhaft, da hier keine DSBI-Enzym kodierenden Sequenzen im Plastom oder Genom verbleiben. Ein entsprechendes Verfahren ist beispielsweise beschrieben bei Segal DJ et al. (1995) Proc Natl Acad Sci USA 92:806-810.
- Das DSBI-Enzym kann als Fusionsprotein mit dem VirE2 oder VirF Protein eines Agrobakterium sowie einer PLS in Pflanzenzellen eingeschleust werden. Entsprechende Verfahren sind beispielsweise für die Cre-Rekombinase beschrieben (Vergunst AC et al. (2000) Science 290:979-982). Diese Ausführungsform ist vorteilhaft, da hier keine DSBI-Enzym kodierenden Sequenzen im Genom verbleiben.

Natürlich sind auch Kombinationen der oben beschriebenen Möglichkeiten denkbar.

- Bevorzugt ist die Expressionskassette für das DSBI-Enzym auf der Insertionssequenz oder dem Transformationskonstrukt enthalten.

Bevorzugt wird das DSBI-Enzym gleichzeitig mit oder nach der Einführung der Insertionssequenz in die Plastiden eingebracht bzw. aktiviert. Die Expression bzw. Aktivierung am richtigen Ort und zur richtigen Zeit kann durch verschiedene Ansätze sichergestellt werden:

45

## a) Induzierbare Expression

Die Expression eines DSBI-Enzym kann unter Verwendung eines induzierbaren Promotor, bevorzugt eines chemisch induzierbaren Promotors, gesteuert werden. Dazu kann beispielsweise die für das DSBI-Enzym kodierende Expressionskassette stabil in die plastidäre oder nukleäre DNA einer Masterpflanze transformiert werden. Erfolgt die Transformation in das Kern-  
5 genom, so muss - wie oben beschriebene - die subzelluläre Lokalisation durch geeignete PLS-Transitpeptide sicher-  
10 gestellt werden. Kurz vor oder während der Transformation mit der Insertionssequenz oder dem Transformationskonstrukt wird dann in Abhängigkeit von dem gewählten induzierbaren System  
15 die Expression des DSBI-Enzyms durch Applikation des entsprechenden Induktors eingeschaltet. Dem Fachmann sind verschiedene Verfahren oder Promotoren zur induzierten Expression bekannt. Als Stimulus können chemische Substanzen oder auch physikalische Stimuli wie beispielsweise erhöhte Temperatur oder Verwundung etc. fungieren. Verschiedene  
20 Beispiele sind weiter unten beschrieben.

## b) Induzierbare Aktivität

Das DSBI-Enzym kann bereits in den Plastiden der Master-  
25 pflanze vorliegen, wenn die Aktivität durch geeignete Techniken erst zum gewählten Zeitpunkt etwa durch Zugabe chemischer Substanzen induziert wird. Entsprechende Verfahren sind für sequenzspezifische Rekombinasen beschrieben (Angrand PO et al. (1998) Nucl Acids Res 26(13):3263-3269; Logie C und Stewart AF (1995) Proc Natl Acad Sci USA 92(13):5940-5944;  
30 Imai T et al. (2001) Proc Natl Acad Sci USA 98(1):224-228). Bei diesen Verfahren werden Fusionsproteine aus dem DSBI-Enzym und der Ligandenbindedomäne eines Steroidhormon-  
rezeptors (z.B. des humanen Androgenrezeptors, oder mutierte  
35 Varianten des humanen Estrogenrezeptors wie dort beschrieben) eingesetzt. Die Induktion kann mit Liganden wie beispielsweise Estradiol, Dexamethason, 4-Hydroxytamoxifen oder Raloxifen erfolgen.

40 Manche DSBI-Enzyme sind als Dimer (Homo- oder Heterodimer) aktiv (I-CreI bildet ein Homodimer; I-PpoI bildet Homodimer, Flick KE et al. (1998) Nature 394: 96-101). Allgemein neigen Enzyme der LAGLIDADG Familie dazu, Homodimere zu bilden, wenn sie je Monomer nur ein LAGLIDADG Motiv enthalten (Jurica MS &  
45 Stoddard BL (1999) Cell Mol Life Sci 55:1304-1326; beispielhaft sei I-CeuI genannt). Eine Dimerisierung kann induzierbar gestaltet werden, indem beispielsweise die natürlichen

27

Dimerisierungsdomänen gegen die Bindungsdomäne eines niedermolekularen Liganden ausgetauscht werden. Zugabe eines dimeren Liganden bewirkt dann Dimerisierung des Fusionsproteins. Entsprechende induzierbare Dimerisierungsverfahren als auch die Herstellung der dimeren Liganden sind beschrieben (Amara JF et al. (1997) Proc Natl Acad Sci USA 94(20): 10618-1623; Muthuswamy SK et al. (1999) Mol Cell Biol 19(10): 6845-6857; Schultz LW und Clardy J (1998) Bioorg Med Chem Lett 8(1):1-6; Keenan T et al. (1998) Bioorg Med Chem. 6(8):1309-1335).

c) Cotransfektion

Bevorzugt wird die Expressionskassette kodierend für das DSBI-Enzym gleichzeitig mit der Insertionssequenz in die Plastiden eingebracht. Dabei können die Expressionskassette für das DSBI-Enzym und die Insertionssequenz auf einem DNA-Molekül oder auf zwei getrennten vorliegen. Bevorzugt liegen die beiden Sequenzen auf einem DNA-Molekül zusammen vor, so dass die Expressionskassette in dem die Insertionssequenz umfassenden Transformationskonstrukt enthalten ist.

Dabei wird in einer besonders bevorzugten Ausführungsform, nachdem homoplastome Pflanzen regeneriert wurden, die Sequenz kodierend für das DSBI-Enzym aus dem Genom der transformierten Plastiden wieder entfernt. Dem Fachmann sind dazu verschiedene Verfahren bekannt, die weiter unten im Detail beschrieben sind.

30 Manche der oben beschriebenen DSBI-Enzyme (insbesondere Homing Endonukleasen) können in dem bevorzugt genutzten Zwischenwirt *E. coli* Erkennungssequenzen aufweisen. Da zudem manche Expressionskassetten zur Expression in Plastiden auch in *E. coli* funktionell sind, wird bevorzugt die Expression des DSBI-Enzyms 35 in *E. coli* auf verschiedene dem Fachmann geläufige Art verhindert, um während der Amplifikation der Expressionskassette etwaige nachteilige Effekte auf *E. coli* zu vermeiden. So können beispielsweise mehrere, aufeinander folgende für *E. coli* seltene Codons (z.B. die für Arginin kodierenden Kodons AGA und AGG) in 40 den Leserahmen des DSBI-Enzyms einfügt werden. Dies verhindert die Expression in *E. coli*, ermöglicht jedoch weiterhin - aufgrund der unterschiedlichen Kodonnutzung - eine Expression in den Plastiden. Alternativ können Promotoren genutzt werden, die nicht in *E. coli* wohl aber in den Plastiden höherer Pflanzen aktiv sind 45 (z.B. Promotoren der nuklear kodierten plastidären RNA-Polymerasen [NEP-Promotoren; s.u.]). Eine bevorzugte Methode ist die Verwendung von Promotoren, die weder von den Plastiden noch

von *E. coli* erkannt werden (z.B. virale oder Bakteriophagen Promotoren), die erst durch das gleichzeitige Vorhandensein der entsprechenden viralen/Bakteriophagen RNA Polymerase funktionell werden. Entsprechende Verfahren sind dem Fachmann bekannt und 5 unten beschrieben. Ferner wäre es denkbar, die entsprechenden DSB-Erkennungssequenzen in *E. coli* zu zerstören oder einen anderen Wirt zu benutzen, der keine DSB-Erkennungssequenzen für das jeweilige DSBI-Enzym besitzt. Des weiteren ist es denkbar und vorteilhaft, die kodierende Region des DSBI-Enzyms zur 10 Amplifikation in *E. coli* promotorlos vorliegen zu haben. In diesem Fall ist die für das DSBI-Enzym kodierende Sequenz bevorzugt auf einem Plasmid enthalten, welches in das plastidäre Genom der zu transformierenden Pflanze integrieren kann. Dabei kann der Integrationsort derart gewählt werden, dass das Gen kodierend 15 für das DSBI-Enzym unter die Kontrolle eines natürlicherweise oder künstlich eingefügt im Plastom vorhandenen Promotors gelangt und es daher in den Plastiden zur Expression des DSBI-Enzyms kommt. In einer weiteren bevorzugten Ausführungsform wird dafür gesorgt, dass das Gen kodierend für das DSBI-Enzym wieder aus dem 20 Plastom deletiert werden kann (siehe unten). Darüber hinaus ist es möglich, eine Verknüpfung zwischen einem Promotor und einem DSBI-Enzym zu erstellen, indem ein solcher in vitro mittels dem Fachmann geläufige PCR Techniken stromaufwärts des offenen Lese-rahmens angefügt wird. Das PCR Produkt kann dann genutzt werden, 25 um es in die pflanzlichen Plastiden einzubringen. Außerdem können nicht-funktionelle Teile einer Expressionskassette für ein DSBI-Enzym erstellt und in *E. coli* amplifiziert werden, wenn diese nach dem Einbringen in pflanzliche Plastiden miteinander rekombinieren (beispielsweise durch homologe Rekombination in überlappenden Bereichen der nicht-funktionellen Anteile der Expressionskassette) und so eine funktionelle Expressionskassette ergeben.

"Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen" (infolge "DSB-Erkennungssequenz" für Doppelstrangbruch-35 Erkennungssequenz) meint allgemein solche Sequenzen, die unter den Bedingungen in den Plastiden der jeweils verwendeten pflanzlichen Zelle oder Pflanze die Erkennung und Spaltung durch ein DSBI-Enzym erlauben. Besonders bevorzugt sind DSB-Erkennungssequenzen für Homing-Endonukleasen, die natürlicherweise in Mitochondrien oder Kern anderer Organismen kodiert sind. Es können auch DSB-Erkennungssequenzen der Homing-Endonukleasen genutzt werden, die aus Plastiden (beispielsweise von Grünalgen) stammen. Bevorzugt ist die DSB-Erkennungssequenz singulär in der plastidären DNA, d.h. ein Doppelstrangbruch wird nur an der so vordefinierten Stelle erzeugt. Es sind jedoch auch Fälle denkbar, bei denen mehr als eine DSB-Erkennungssequenz im Plastom vorhanden ist. Dies ist insbesondere der Fall, wenn die DSB-

29

Erkennungssequenz in duplizierten Genen (z.B. in invertierten "Repeats") lokalisiert ist. Im letzteren Fall liegen mehr als eine identische DSB-Erkennungssequenz vor, ihr Kontext ist jedoch identisch, so dass auch hier eine gezielte Insertion erfolgt. Es 5 ist sogar bevorzugt, dass die Integration in alle Kopien erfolgt, so dass auch ein Schnitt in allen Kopien erforderlich ist. DSB-Erkennungssequenzen, die zwar mehr als einmal in einem Plastom auftreten, jedoch im gleichen plastom Kontext lokalisiert sind (z.B. in "Repeats" oder Genduplikationen) sind im Rahmen 10 dieser Erfindung unter dem Begriff der singulären DSB-Erkennungssequenzen subsumiert.

Bevorzugt ist die eingesetzte Pflanze oder von dieser abgeleitete Zelle bezüglich der DSB-Erkennungssequenz überwiegend homoplastom 15 oder homotransplastom, d.h. dass die überwiegende Anzahl der in einem Plastid enthaltenen plastidären DNA-Moleküle diese DSB-Erkennungssequenz aufweisen. Solche Pflanzen werden auch als Masterpflanzen im Rahmen dieser Erfindung bezeichnet.

20 Prinzipiell können zwei Arten von DSB-Erkennungssequenzen genutzt werden:

a) Natürliche, endogene DSB-Erkennungssequenzen

25 Wie im Rahmen dieser Erfindung gezeigt werden konnte, enthalten die Plastome höherer Pflanzen verschiedene Sequenzen, die als Erkennungssequenzen für DSBI-Enzyme (beispielsweise Homing-Endonukleasen) fungieren können, auch wenn für höhere Pflanzen bislang keine solchen Endonukleasen nachgewiesen 30 werden konnten. Solche DSB-Erkennungssequenzen können mittels Durchmusterung der plastidären DNA-Sequenz unter Verwendung der bekannten DSB-Erkennungssequenzen (beispielsweise der in Tabelle 2 beschriebenen) identifiziert werden. Das plastidäre Genom verschiedener Pflanzen ist bekannt 35 ([http://megasun.bch.umontreal.ca/ogmp/projects/other/cp\\_list.html](http://megasun.bch.umontreal.ca/ogmp/projects/other/cp_list.html)). Berichtet sind die Sequenzen der Plastome von

- *Arabidopsis thaliana* (Sato S et al. (1999) DNA Res. 6 (5):283-290) (GenBank Acc.-No.: AP000423; NCBI Acc.-No. 40 NC\_000932)
- *Epifagus virginiana* (Beechdrops; Wolfe KH et al. (1992) J Mol Evol 35(4):304-317; NCBI Acc.-No.: NC\_001568; GenBank Acc.-No.: M81884).

45

30

- *Lotus japonicus* (Kato T et al. (2000) *DNA Res* 7(6):323-330; NCBI Acc.-No.: NC\_002694; GenBank Acc.-No.: AP002983)
- 5 - *Oryza sativa* (Reis) (Hiratsuka J et al. (1989) *Mol Gen Genet* 217(2-3):185-194; NCBI Acc.-No.: NC\_001320; GenBank Acc.-No.: X15901).
- 10 - *Marchantia polymorpha* (Liverwort; Ohyama K et al. (1988) *J Mol Biol* 203(2):281-298; Yamano Y et al. (1984) *Nucl Acids Res* 12(11):4621-4624; GenBank Acc.-No.: X04465 und Y00686; NCBI Acc.-No.: NC\_001319)
- 15 - *Nicotiana tabacum* (Tabak) (GenBank Acc.-No.: Z00044 und S54304; NCBI Acc.-No.: NC\_001879; Shinozaki K et al. (1986) *EMBO J* 5:2043-2049)
- 20 - *Oenothera elata* ssp. *hookeri* (Monterey evening primrose; GenBank Acc.-No.: AJ271079; NCBI Acc.-No.: NC\_002693; Hupfer H et al. (2000) *Mol Gen Genet* 263(4):581-585)
- 25 - *Medicago truncatula* (Gen Bank Acc.-No.: AC093544)
- *Pinus thunbergii* (black pine; Tsudzuki J et al. (1994) *Curr Genet* 26(2):153-158; NCBI Acc.-No.: NC\_001631; GenBank Acc.-No.: D17510)
- 30 - *Spinacia oleracea* (GenBank Acc.-No.: AJ400848 J01442 M12028 M16873 M16878 M27308 M55297 X00795 X00797 X01724 X04131 X04185 X05916 X06871)
- *Triticum aestivum* (Weizen; GenBank Acc.-No.: AB042240; NCBI Acc.-No.: NC\_002762) und
- 35 - *Zea mays* (GenBank Acc.-No.: X86563; NCBI Acc.-No.: NC\_001666)

Darüber hinaus können weitere Plastome sequenziert werden, um auch dort DSB-Erkennungsstellen zu identifizieren. Es ist im allgemeinen ausreichend, hoch konservierte Regionen aus dem Plastom durch den Fachmann geläufige PCR-Methoden zu isolieren und nur diese zu sequenzieren.

Ferner können natürliche, endogene DSB-Erkennungsstellen experimentell ermittelt werden, indem man beispielsweise die plastidäre DNA isoliert (z.B. nach Mariac P et al. (2000) *BioTechniques* 28:110-113), die zu betrachteten Fragmenten des Plastidengenoms mittels PCR amplifiziert oder synthetische

## 31.

Fragmente nutzt und mit dem jeweiligen DSBI-Enzym eine Restriktionsanalyse durchführt. Diese wird bevorzugt unter Bedingungen durchgeführt, wie sie im Plastid einer höheren Pflanze vorliegen.

5

Die im Rahmen dieser Erfindung identifizierten und in Tabelle 1 beschriebenen endogenen DSB-Erkennungssequenzen für natürliche Homing-Endonukleasen liegen darüberhinaus in den konservierten Regionen des Plastoms, so dass - insbesondere unter Berücksichtigung der gegebenen Variabilität der entsprechend genannten Homing-Endonukleasen bezüglich ihrer jeweiligen Erkennungssequenzen - davon ausgegangen werden kann, dass diese Erkennungssequenzen nahezu universell in allen Plastomen höherer Pflanzen vorkommen. Hinter den in Tabelle 1 angegebenen Positionen verbirgt sich jeweils die angegebene Sequenz und die reverskomplimentäre, da alle in Tabelle 1 angegebenen Erkennungsregionen in den "Inverted Repeats" des Plastidengenoms lokalisiert sind. Von den in Tabelle 1 genannten Homing-Endonukleasen sind besonders bevorzugt I-CpaI, I-CeuI, I-ChuI, I-CpaII und I-CreI.

10

15

20

25

30

35

40

Die so identifizierten Erkennungssequenzen können für die Insertion von Fremd-DNA genutzt werden, indem durch Einführung des entsprechenden DSBI-Enzyms ein Doppelstrangbruch erzeugt wird. Sollte die DSB-Erkennungssequenz in einer hochkonservierten Region innerhalb eines Gens des Organellen-Genoms liegen, so wird die Fremd-DNA bevorzugt in Form eines selbstspleißenden Introns insertiert, was die Rekonstitution der mRNA des betroffenen Gens ermöglicht (s.u.).

Dem Fachmann sind ferner Verfahren bekannt, nach denen eine beliebige endogene Sequenz als Erkennungssequenz für chimäre, mutierte oder künstliche Endonukleasen fungieren kann, indem deren DNA-Erkennungsregion - beispielsweise durch Modifizieren einer an eine Endonuklease-Domäne fusionierten Zink-Finger-Domäne oder Modifizieren der RNA-Sequenz eines Gruppe II Intron-RNA-Protein Komplexe - gezielt verändert wird (s.o.; WO 96/06166, Bibikova M et al. (2001) Mol Cell Biol 21:289-297).

45

5

10

15

20

25

30

35

40

45

DSBI-Enzym	Publizierte DSB-Erkennungssequenz	Sequenz im Plastom	Position im Tabakplastom Acc. 200044	Position im Weizenplastom Acc. AB042240	Position im Reisplastom Acc. X159019	Position im Maisplastom Acc. X86563	Position im Arabidopsis-plastom Acc. AP000423
I-DmoI	ATGGCGCCGGAACCTACCCGGCAAGGCAT	GTCGGGTGGAACTT ACCCGACAAGGAAT	C(108281-108310) 1343116-134345	118010-118039 c(96855- 96884)	117846-117875 c(97243- 97272)	121617-121646 c(101091- 101120)	131977-132006 c(106643- 106672)
I-CpaI	CGATCCTAAGTAGTCGAAATTCA	CGGTCTTAAGGTAGCG AAATTCC	108263-108285 c(134341-134363)	96837-96859 c(118035- 118057)	97224-97246 c(117871- 117893)	101073-101095 c(121642- 121664)	106625-106647 c(132002- 132024),
I-CeuI	CGTAACATAAACGGTCTTAAGGTAGCGAA	CGTAACATAAACGGT CTTAAGGTAGCGAA	C(134346-134374) 108252-108280	c(118040- 118068)	C(117876- 117904)	101062-101090 c(121647- 121675)	132035),
I-ChuI	GAAGGTTTGGCACCTCGATGTCGGCTCATC	GAAGGTTTGGCACCT CGATGTCGGCTCATC	108832-108861 c(133765-133794)	97405-97434 c(117460- 117489)	97792-97821 c(117297- 117326)	101641-101670 c(121067- 121096)	107194-107223 c(131426- 131455)
I-CpaII	CCCCGCTTAACCTCTGTGCGAG	ATGGGCTTAACCTCTGT CCAG	C(139398-139417) 103209-103228	c(123374- 123393)	C(123351- 123370)	C(127108- 127127)	C(137169- 137188)
I-CreI	CTGGGTTTACAAACGTCTGAGACGG	CTGGGTTTACAAACGT CTGAGACAGTTGG	108925-108954 c(133672-133701)	97498-97527 c(117367- 117396)	97785-97914 c(117204- 117233)	95610-95629 101461-101480	107287-107316 abweichende Sequenz: ctgggttccagaac gtcggtgagacgtt cgg c(120975- 121003), 101734-101762
I-SceI	TACCCCTGTATCCCTAGGCTAACT	C(108804-108781) 133822-133845	C(97377- 97354)	C(97741- 97764)	C(101590- 101613)	C(107143- 107166)	
			117516-117540	117354-117377	121124-121147	131483-131506	

Tab 1: Bevorzugte, endogene Schnittstellen in den plastidären Genomen von Tabak, Weizen, Reis, Mais und Arabidopsis. c=komplementär. AccNo: GenBank Accession-Number (<http://www.ncbi.nlm.nih.gov/>).

33

5 Darüber hinaus liegen singuläre Schnittstellen von Restriktionsendonukleasen im Plastidengenom vor. Diese befinden sich jedoch meist in weniger hoch konservierten Regionen und sind daher nicht unbedingt universell in allen Pflanzenarten nutzbar. Beispielhaft sind zu nennen:

- 10 a) Das Enzym SfiI hat im Plastidengenom von Arabidopsis (GenBank Acc.-No.: AP000423) an der Position 40846-40858 mit der Sequenz GGCCTTATGGCC eine singuläre Erkennungsstelle.
- 15 b) Im Plastidengenom von Mais (GenBank Acc.-No.: X86563) gibt es eine singuläre Schnittstelle für das Enzym AscI an der Position 42130-42137 mit der Sequenz GGCGCGCC.
- 20 c) Im Plastidengenom von Reis (GenBank Acc.-No.: X159019) gibt es eine singuläre Schnittstelle für das Enzym SgfI an der Position 77309-77316 mit der Sequenz GCGATCGC sowie für das Enzym AscI an der Position 39776-39783 mit der Sequenz GGCGCGCC.
- 25 d) Im Plastidengenom von Tabak (Accession Z00044) gibt es je eine singuläre Schnittstelle für das Enzym SfiI an der Position 42475-42487 mit der Sequenz GGCCTTATGGCC, für das Enzym SgrI an der Position 78522-78529 mit der Sequenz CACCGGCG, sowie für das Enzym PmeI an der Position 120895-120902 mit der Sequenz GTTTAAC.
- 30 e) Im Plastidengenom von Weizen (Accession AB042240) gibt es je eine singuläre Schnittstelle für das Enzym PmeI an der Position 59331-59338 mit der Sequenz GTTTAAC, für die Enzyme NarI, KanI, EheI und BbeI eine singuläre Schnittstelle an der Position 41438-41443 mit der Erkennungssequenz GGCGCC, sowie für das Enzym SfiI eine Erkennungsregion an der Position 112656-112668 mit der Sequenz GGCCCAGGGGGCC.

35 All diese Pflanzen mit endogenen, natürlichen DSB-Erkennungssequenzen stellen quasi natürlich vorkommende "Masterpflanzen" dar. Bei ihnen ist die DSB-Erkennungssequenz natürlicherweise homoplastom vorhanden. Dies erübriggt die Einführung und Selektion künstlicher DSB-Erkennungssequenzen.

b) Künstlich eingeführte DSB-Erkennungssequenzen

Dem Fachmann ist bewusst, dass die in eine Masterpflanze eingebrachte Erkennungsregion für ein selten schneidendes Enzym nicht natürlich sein muss. Prinzipiell kann jede Erkennungssequenz eines beliebigen DSBI-Enzyms an jede beliebige Stelle der plastidären DNA insertiert werden. Die Herstellung erfolgt bevorzugt unter Verwendung eines Konstrukt zur Insertion der DSB-Erkennungssequenz (infolge DSBE-Konstrukt). Bevorzugt umfasst das DSBE-Konstrukt einen Selektionsmarker, um die zur Erzeugung entsprechender Masterpflanzen erforderliche Selektion transplastomer Pflanzen mit der erfolgreich insertierten DSB-Erkennungssequenz zu erleichtern. Dem Fachmann sind verschiedene Selektionsmarker bekannt, die eine Selektion von Plastiden ermöglichen (s.u.). Bevorzugt sind aadA, nptII oder BADH, wobei aadA besonders bevorzugt ist. Die Selektion erfolgt beispielsweise mit Hilfe des dem Fachmann bekannten "segregation and sorting" Prozess (beispielhaft unter Beispiel 4 beschrieben). Der Selektionsmarker ist bevorzugt so konstruiert, dass eine nachfolgende Deletion aus dem Plastom ermöglicht wird. Entsprechende Verfahren sind dem Fachmann bekannt und unten beschrieben.

Bevorzugt wird so zunächst eine bezüglich der insertierten DSB-Erkennungssequenz homotransplastome Pflanze erzeugt, die eine DSB-Erkennungssequenz in allen oder der überwiegenden Anzahl der Plastiden der betrachteten Pflanze besitzen. Solche Pflanzen können vorteilhaft als Masterpflanzen eingesetzt werden.

Das DSBE-Konstrukt kann neben dem Selektionsmarker weitere Sequenzen enthalten. Diese können beispielsweise weitere regulatorische Elemente für die Expression der infolge einzuführenden Insertionssequenzen enthalten. Der im Rahmen des Konstruktes zur Insertion der DSB-Erkennungssequenz eingeführte Selektionsmarker wird in einer bevorzugten Ausführungsform nach Erhalt der homoplastomen Masterpflanze durch den Fachmann bekannte Verfahren deletiert (s.u.).

In einer bevorzugten Ausführungsform umfasst das DSBE-Konstrukt zur Ermöglichung einer ortsspezifischen Insertion an mindestens einer, bevorzugt an beiden Seiten der DSB-Erkennungssequenz weitere flankierende Sequenzen, die eine ausreichende Länge und Homologie zu entsprechenden Zielsequenzen im Plastom haben, um eine ortsspezifische Insertion mittels homologer Rekombination zu gewährleisten.

Aufgrund der Vielzahl der im Stand der Technik beschriebenen DSBI-Enzyme mit definierten Erkennungssequenzen, ist es möglich und bevorzugt, Master-Pflanzen zu erzeugen, die mehrere verschiedene singuläre DSB-Erkennungssequenzen in ihr plastidäres Genom eingebaut haben.

Beispielhaft aber nicht einschränkend seien dabei in nachfolgender Tabelle 2 die Erkennungssequenzen für die jeweiligen aufgeführten DSBI-Enzyme genannt.

10

Tabelle 2: Erkennungssequenzen und Herkunftsorganismus von DSBI-Enzymen ("^" gibt innerhalb einer Erkennungssequenz die Schnittstelle des DSBI-Enzyms an.)

15	DSBI-Enzym	Herkunfts-organismus	Erkennungssequenz
	I-AniI	Aspergillus nidulans	5' -TTGAGGGAGGTT^TCTCTGAAATAAANNNNNNNNNNNNN 3' -AACTCCTCCAAAGAGACATTATNNNNNNNNNNNNNN
20	I-CvuI	Chlorella vulgaris	5' -CTGGGTTCAAAACGTCGTGA^GACAGTTGG 3' -GACCCAAGTTTGCAG^CACTCTGTCAAACC
	I-CsmI	Chlamydomonas smithii	5' -GTACTAGCATGGGTCAAATGTCTTCTGG
	I-CmoI	Chlamydomonas moewusii	5' -TCGTAGCAGCT^CACGGTT 3' -AGCATCG^TCGAGTGCCAA
25	I-CreI	Chlamydomonas reinhardtii	5' -CTGGGTTCAAAACGTCGTGA^GACAGTTGG 3' -GACCCAAGTTTGCAG^CACTCTGTCAAACC
	I-ChuI	Chlamydomonas humicola	5' -GAAGGTTTGGCACCTCG^ATGTCGGCTCATC 3' -CTCCAAACCGTG^GAGCTACAGCCGAGTAG
	I-CpaI	Chlamydomonas pallidostigmatica	5' -CGATCCTAACGGTAGCGAA^ATTCA 3' -CCTAGGATTCCATC^GCTTTAAGT
30	I-CpaII	Chlamydomonas pallidostigmatica	5' -CCCGGCTAACTC^TGTGCCAG 3' -GGGCCGAT^TGAGACACGGTC
	I-CeuI	Chlamydomonas eugametos	5' -CGTAACTATAACGGTCCTAA^GGTAGCGAA 3' -GCATTGATATTGCCAG^GATTCCATCGCTT
35	I-DmoI	Desulfurococcus mobilis	5' -ATGCCCTGCCGGTAA^GTTCCGGCGGCAT 3' -TACGGAACGGCC^CATTCAAGGCCGCGTA
	I-SceI	Saccharomyces cerevisiae	5' -AGTTACGCTAGGGATAA^CAGGGTAATATAG 3' -TCAATGCGATCCC^TATTGTCCTTATATC 5' -TAGGGATAA^CAGGGTAAT 3' -ATCCC^TATTGTCCTTATTA ("Core"-Sequenz)
40	I-SceII	Saccharomyces cerevisiae	5' -TTTGATTCTTGGTCACCC^TGAAGTATA 3' -AAAACTAAGAAACCAG^TGGGACTTCATAT
	I-SceIII	Saccharomyces cerevisiae	5' -ATGGAGGTTTGGTAAC^TATTATTC 3' -TAACCTCCAAACC^ATTGATAAAATGG
45	I-SceIV	Saccharomyces cerevisiae	5' -TCTTTCTCTGATTA^GCCCTAATCTACG 3' -AGAAAAGAGAAC^TAATCGGGATTAGATGC
	I-SceV	Saccharomyces cerevisiae	5' -ATAAATTCT^TCTTAGTAATGCC 3' -TTATTAAGAAGAACATTTA^CGG

DSBI-Enzym	Herkunfts-organismus	Erkennungssequenz
I-SceVI	Saccharomyces cerevisiae	5'-CTTATTTAATG^TTTAGTAGTGTGG 3'-CAATAAATTACAAAATCATCA^ACC
5 I-SceVII	Saccharomyces cerevisiae	5'-TGTACACATTGAGGTGCACTAGTTATTAC
PI-SceI	Saccharomyces cerevisiae	5'-ATCTATGTCGGGTGC^GGAGAAAGAGGTAAT 3'-TAGATACAGCC^CACGCCCTTTCTCCATTAA
F-SceI	Saccharomyces cerevisiae	5'-GATGCTGTAGGC^ATAGGCTTGGTT 3'-CTACGACA^TCCGTATCCGAACCAA
10 F-SceII	Saccharomyces cerevisiae	5'-CTTTCCGCAACA^GTAAAATT 3'-GAAAGCGC^TTGTCAATTAA
I-LlaI	Lactococcus lactis	5'-CACATCCATAAAC^CATATCATTTC 3'-GTGTAGGTATTGGTATAGTAA^AAA
15 I-MsoI	Monomastix species	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTGG 3'-GACCCAAGTTTGAG^CACTCTGTCAAACC
I-NanI	Naegleria andersoni	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACCC^ACGGTCGTGGCG
I-NitI	Naegleria italica	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACCC^ACGGTCGTGGCG
20 I-NjaI	Naegleria jamiesoni	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACCC^ACGGTCGTGGCG
I-PakI	Pseudendoclo-nium akinetum	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTGG 3'-GACCCAAGTTTGAG^CACTCTGTCAAACC
I-PorI	Pyrobaculum organotrophum	5'-GCGAGCCCCGTAAGGGT^GTGTACGGG 3'-CGCTCGGGCATT^CCACACATGCC
25 I-PpoI	Physarum polycephalum	5'-TAACTATGACTCTCTTAA^GGTAGCCAAAT 3'-ATTGATACTGAGAG^AATTCCATCGTTTA "Core Sequenz": CTCTCTTAA^GGTAGC GAGAG^AATTCCATCG
30 I-ScaI	Saccharomyces capensis	5'-TGTACACATTGAGGTGCACT^AGTTATTAC 3'-ACAGTGTAACTCCAC^GTGATCAATAATG
I-Ssp6803I	Synechocystis species	5'-GTCGGGCT^CATAACCCGAA 3'-CAGCCCGAGTA^TTGGGCTT
PI-PfuI	Pyrococcus furiosus Vc1	5'-GAAGATGGGAGGAGGG^ACCGGACTCAACTT 3'-CTCTACCCCTCC^TCCCTGGCCTGAGTTGAA
35 PI-PfuII	Pyrococcus furiosus Vc1	5'-ACGAATCCATGTGGAGA^AGAGCCTCTATA 3'-TGCTTAGGTACAC^CTCTCTCGGAGATAT
PI-PkoI	Pyrococcus kodakaraensis KOD1	5'-GATTTTAGAT^CCCTGTACC 3'-CTAAAA^TCTAGGGACATGG
40 PI-PkoII	Pyrococcus kodakaraensis KOD1	5'-CAGTACTACG^GTTAC 3'-GTCATG^ATGCCAATG
PI-PspI	Pyrococcus sp.	5'-AAAATCCTGGCAAACAGCTATTAT^GGGTAT 3'-TTTTAGGACCGTTGTGCGAT^AATACCCATA
45 PI-Tful	Thermococcus fumicola ST557	5'-TAGATTTAGGT^CGCTATATCCTTCC 3'-ATCTAAAA^TCCAGCGATATAGGAAGG

DSBI-Enzym	Herkunfts-organismus	Erkennungssequenz
PI-TfuII	Thermococcus fumicolans ST557	5'-TAYGCNGAYACN^GACGGYTTYT 3'-ATRCGNCT^RTGNCTGCCRAARA
PI-ThyI	Thermococcus hydrothermialis	5'-TAYGCNGAYACN^GACGGYTTYT 3'-ATRCGNCT^RTGNCTGCCRAARA
PI-TliI	Thermococcus litoralis	5'-TAYGCNGAYACNGACGG^YTTYT 3'-ATRCGNCTRGN^TGCCRARA
PI-TliII	Thermococcus litoralis	5'-AAATTGCTTGCAAACAGCTATTACGGCTAT
I-TevI	Bacteriophage T4	5'-AGTGGTATCAAC^GCTCAGTAGATG 3'-TCACCATAGT^TGCAGTCATCTAC
I-TevII	Bacteriophage T4	5'-GCTTATGAGTATGAAGTGAACACGT^TATTC 3'-CGAATACTCATACACTTCACTTGTG^CAATAAG
F-TevI	Bacteriophage T4	5'-GAAAACAAGA^AATGTTTAGTAAANNNNNNNNNNNNN 3'-CTTGTGTTCTTACAATCATTNNNNNNNNNNNN^
F-TevII	Bacteriophage T4	5'-TTAACCTCGCTTC^AGATATGGCAACTG 3'-AAATTAGGAGCGA^AGCTATACC GTTGAC

20

Dabei sind auch Abweichungen (Degenerationen) der Erkennungssequenz umfasst, die dennoch eine Erkennung und Spaltung durch das jeweilige DSBI-Enzym ermöglichen. Derartige Abweichungen - auch in Zusammenhang mit unterschiedlichen Rahmenbedingungen wie beispielsweise Calcium oder Magnesium-Konzentration - sind beschrieben (Argast GM et al. (1998) J Mol Biol 280: 345-353). Ferner sind Kernsequenzen ("Core"-Sequenzen) dieser Erkennungssequenzen umfasst. Es ist bekannt, dass auch die inneren Anteile der Erkennungssequenzen für einen induzierten Doppelstrangbruch genügen und dass die äußeren nicht unbedingt relevant sind, jedoch die Effizienz der Spaltung mitbestimmen können. So kann beispielsweise für I-SceI eine 18bp-"Core"-Sequenz definiert werden. Der Begriff der DSB-Erkennungssequenz umfasst insofern auch alle wesentlichen gleichen Erkennungssequenzen. Im wesentlichen gleiche Erkennungssequenzen meint solche Erkennungssequenzen, die zwar Abweichungen von der für das jeweilige Enzym als optimal gefundenen Erkennungssequenz aufweisen, jedoch eine Spaltung durch dasselbe noch erlauben.

40 Es sind verschiedene Orte der Lokalisation (bei bereits vorhandenen endogenen DSB-Erkennungsequenzen) bzw. Integration (bei künstlich generierten DSB-Erkennungssequenzen) für die DSB-Erkennungssequenz möglich. Beispielhaft seien zu nennen:

45

## 38.

- a) Lokalisation (Integration) in einer transkriptionell stillen Region

5 Lokalisation (Integration) der DSB-Erkennungssequenz in einer transkriptionell stillen Region des Plastidengenoms (intergenicische Region) ist die bevorzugte Ausführungsform. Eine Störung der plastidären Funktionen kann so weitgehend ausgeschlossen werden. Hierbei ist zu beachten, dass für eine Expression gegebenenfalls entsprechende regulatorische Elemente wie Promotoren etc. mit eingebracht werden müssen.

- 10 b) Lokalisation (Integration) in eine transkriptionell aktive, aber nicht-kodierende (intervistronische) Region

15 Diese Lokalisation (Integration) hat den Vorteil, dass dadurch die einzubringende Insertionssequenz letztendlich in einem plastidären Operon kodiert ist und Promotor(en) bzw. Terminator(en) nicht gesondert mit eingebracht werden müssen, sondern die endogen an diesem Locus vorhandenen ausgenutzt werden können (aber nicht müssen). Es sollten in einem solchen Fall lediglich Ribosomenbindestellen in geeignetem Abstand stromaufwärts der kodierenden Region der einzubringenden Fremd-Gene vorhanden sein.

20 25 Es ist jedoch auch denkbar, dass eine intergenische Region nicht vollständig transkriptionell still ist, weil beispielsweise eine nur ineffiziente Termination der Transkription von einem benachbarten Gen oder Operon erfolgt.

- 30 c) Lokalisation (Integration) in eine transkriptionell aktive, kodierende Regionen.

35 Die unter a) und b) beschriebene Lokalisation (Integration) der DSB-Erkennungssequenz an einem nicht kodierenden Locus, hat den Vorteil, dass die Insertion der Fremd-DNA die Funktion des Plastidengenoms mit hoher Wahrscheinlichkeit nicht beeinflusst. Nicht-kodierende Bereiche sind jedoch weniger konserviert als kodierende. Um ein möglichst universelles Verfahren zu haben, dass in vielen Pflanzenarten funktioniert, ist in einer besonders bevorzugten Ausführungsform, 40 die DSB-Erkennungssequenz (und infolge die Insertionssequenz) in der kodierende Sequenz eines bestehenden Gens lokalisiert. Die Zerstörung der Genfunktion durch die Einführung der DSB-Erkennungssequenz (bei einer künstlich generierten DSB- 45 Erkennungssequenz) oder die Einführung der Insertionssequenz wird auf erfinderische Weise dadurch verhindert, dass die DSB-Erkennungssequenz bzw. die Insertionssequenz in einer

bevorzugten Variante dieser Ausführungsform im Rahmen eines Introns eingebracht wird. Auf diese Weise wird die vollständige, kodierende mRNA durch Spleißen der pre-RNA des Gens am Integrationsort wieder generiert.

5

Nicht natürlicherweise in der plastidären DNA vorkommende DSB-Erkennungssequenzen können auf verschiedene Arten in die plastidäre DNA eingeführt werden. Beispielhaft seien zu nennen:

10 a) Integration mittels Doppel-Crossover

Bevorzugt wird die Integration in das Plastidengenom mit Hilfe der oben beschriebenen, dem Fachmann allgemein bekannten Verfahren (Doppel-Crossover) durchgeführt.

15

b) Integration unter Verwendung natürlicher, endogener DSB-Erkennungssequenzen

c) Integration unter Verwendung von Rekombinasen und entsprechenden Erkennungssequenzen.

20

Auch wenn der Aufwand zur Insertion einer künstlichen DSB-Erkennungssequenz in die plastidäre DNA relativ hoch ist und im Fall von a) dem der zur Zeit im Stand der Technik beschriebenen 25 Verfahren zur Plastidentransformation entspricht, so muss dieser Aufwand lediglich einmalig betrieben werden. Die erhaltene homoplastome Masterpflanze kann dann für beliebig viele unterschiedliche nachfolgende Transformationen unter Verwendung des erfindungsgemäßen Verfahrens eingesetzt werden, was eine erhebliche Steigerung der Transformationseffizienz ermöglicht: Statt 30 den herkömmlichen Selektionsprozess für eine homotransplastome Pflanze jedes Mal vollständig durchlaufen zu müssen, muss er hier nur einmal realisiert werden.

35 "Deaktivierung der Funktionalität" einer DSB-Erkennungssequenz meint, dass durch Insertion der Insertionssequenz an oder nahe der Position des Doppelstrangbruches eine Zerstörung der DSB-Erkennungssequenz erfolgt, d.h., dass das entsprechende DSBI-Enzym die Region nicht mehr erkennt und dementsprechend dort 40 keinen Doppelstrangbruch mehr induziert.

Aufbau des Transformationskonstruktes mit der Insertionssequenz

Unter Verwendung einer der oben beschriebenen Masterpflanzen bzw. 45 von diesen abgeleiteten Zellen, die eine natürliche und/oder eine künstlich-erzeugte DSB-Erkennungssequenz im Plastom enthalten, wird die Insertionssequenz in besagte DSB-Erkennungssequenz

40.

im Rahmen einer Transformation insertiert. Dies geschieht bei gleichzeitigem Vorhandensein eines DSBI-Enzyms, welches eine der DSB-Erkennungssequenzen im Plastom erkennt.

- 5 In seiner einfachsten Form besteht das Transformationskonstrukt allein aus der Insertionssequenz selber, beispielsweise aus einer Expressionskassette, die die Expression eines bestimmten Gens in den Plastiden gewährleisten soll. Die sequenzspezifische Induktion von Doppelstrangbrüchen ist ausreichend, um die  
10 Platzierung dieser Insertionssequenz an dieser Position zu gewährleisten und so die Deaktivierung der DSB-Erkennungssequenz zu bewirken.

In einer bevorzugten Ausführungsform umfasst die Insertionssequenz mindestens eine zu exprimierende Nukleinsäuresequenz. Um die Expression (Transkription und/oder Translation) zu gewährleisten, sind diese - je nach Ausführungsform und Insertionsort - mit regulatorischen Elementen zu versehen. Erfolgt die Insertion in einen transkriptionell aktiven Lokus so sind - wie oben beschrieben - keine Promotorsequenzen erforderlich. Vorteilhaft werden die zu exprimierenden Sequenzen in jedem Fall mit Ribosomenbindestellen in geeignetem Abstand stromaufwärts des offenen Leserahmens versehen oder besitzen bereits solche von Natur aus. Diese regulatorischen Sequenzen oder Teile derselben können aber auch im Plastom natürlicherweise vorhanden sein oder bereits im ersten Schritt, d.h. bei der Generierung einer nicht-natürlichen Masterpflanze zusammen mit der DSB-Erkennungssequenz in die plastidäre-DNA eingeführt werden.

- 30 Eine Steigerung der Insertionseffizienz und -genauigkeit kann dadurch bewirkt werden, dass die im Transformationskonstrukt enthaltene Insertionssequenz und die DSB-Erkennungssequenz von homologen Sequenzbereichen flankiert werden, die eine homologe Rekombination infolge des induzierten Doppelstrangbruches  
35 gewährleisten. In einer bevorzugten Ausführungsform umfasst die Insertionssequenz flankierende Homologiesequenzen A' bzw. B', wobei die in die plastidäre DNA einzufügende Sequenz zwischen A' und B' liegt. Die DSB-Erkennungssequenz wird von Homologiesequenzen A bzw. B flankiert, wobei die DSB-Erkennungssequenz  
40 zwischen A und B liegt. A und B können natürlichen Ursprungs sein, oder im Rahmen der Insertion von nicht-natürlichen DSB-Erkennungssequenzen eingeführt worden sein. A und A' bzw. B und B' weisen eine ausreichende Länge und ausreichende Homologie zu einander auf, um eine homologe Rekombination zwischen A und A'  
45 bzw. B und B' zu gewährleisten.

41

In einer weiteren Ausführungsform ist die DSB-Erkennungssequenz lediglich von einer Homologiesequenz A flankiert, die eine ausreichende Homologie zu einer die Insertionssequenz einseitig flankierenden Sequenz A' aufweist.

5

"Ausreichende Länge" meint in Bezug auf die Homologiesequenzen bevorzugt Sequenzen von einer Länge von mindestens 20 Basenpaaren, bevorzugt mindestens 50 Basenpaaren, besonders bevorzugt von mindestens 100 Basenpaaren, ganz besonders bevorzugt von 10 mindestens 250 Basenpaaren, am meistens bevorzugt von mindestens 500 Basenpaaren.

"Ausreichende Homologie" meint in Bezug auf die Homologiesequenzen A und A' bzw. B und B' bevorzugt Sequenzen die eine 15 Homologie innerhalb dieser Homologiesequenzen aufweisen von mindestens 70 %, bevorzugt 80 %, vorzugsweise mindestens 90 %, besonders bevorzugt mindestens 95 %, ganz besonders bevorzugt mindestens 99 %, am meisten bevorzugt 100% über eine Länge von mindestens 20 Basenpaaren, bevorzugt mindestens 50 Basenpaaren, 20 besonders bevorzugt von mindestens 100 Basenpaaren, ganz besonders bevorzugt von mindestens 250 Basenpaaren, am meistens bevorzugt von mindestens 500 Basenpaaren.

Unter Homologie zwischen zwei Nukleinsäuren wird die Identität 25 der Nukleinsäuresequenz über die jeweils gesamte Sequenzlänge verstanden, die durch Vergleich mit Hilfe des Programmalgorithmus GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA) unter Einstellung folgender Parameter berechnet wird:

30

Gap Weight: 12

Length Weight: 4

Average Match: 2,912

Average Mismatch:-2,003

35 Da die homologe Rekombination durch den induzierten Doppelstrangbruch gefördert wird, sind die Anforderungen an Länge und Homologie der Sequenzen deutlich geringer als beispielsweise bei einer konventionellen homologen Rekombination. Die homologen Bereiche können dabei auch deutlich kleiner als 250 bp sein. Die 40 Verwendung von Homologiesequenzen, hat den Vorteil, dass wenn A' und B' unterschiedlich sind, oder lediglich eine Homologiesequenz A' verwendet wird, eine gerichtete Insertion der Insertionssequenz in die plastidäre DNA erfolgen kann.

45 Bevorzugt umfasst das Transformationskonstrukt oder die Insertionssequenz einen Selektionsmarker, der eine Selektion transplastomer Plastiden ermöglicht (s.u.), besonders bevorzugt

## 42

aadA, BADH oder einen "binding type"-Marker. Der Selektionsmarker ist bevorzugt so konstruiert, dass eine nachfolgende Deletion aus dem Plastom ermöglicht wird. Entsprechende Verfahren sind dem Fachmann bekannt und unten beschrieben.

5

Die Insertionssequenz oder das Transformationskonstrukt hat bevorzugt die Struktur und Sequenz eines Introns. In der Regel werden dazu natürlicherweise vorkommende Introns so modifiziert, dass sie den Erfordernissen des erfindungsgemäßen Verfahrens

10 genügen. Derartige artifizielle Introns sind insbesondere bevorzugt, wenn die Insertion in eine transkriptionell aktive oder gar kodierende Region - beispielsweise unter Nutzung einer natürlichen, endogenen DSB-Erkennungssequenz - erfolgen soll. Bevorzugt erfolgt die Insertion derart, dass die insertierte Sequenz 15 durch Spleißen der pre-mRNA restlos entfernt wird. Die herausgespleißte RNA (also das artifizielle Intron) stellt dann die mRNA beispielsweise für die Translation auf ihr kodierter Proteine dar. Dieses Verfahren hat weitere Vorteile:

20 - Die genutzten Introns weisen eine ausgeprägte Sekundäralfaltung auf, so dass sich eine relativ stabile RNA ergibt. Dadurch können die in dem Intron kodierten Gene von Interesse besonders hoch exprimiert werden, wie es beispielsweise in E. coli demonstriert wurde (Chan KYW et al. (1988) Gene 73:295-304).

25 - Wenn das Intron in ein Gen hinein integriert wird, unterliegt die Transkription des Introns der regulatorischen Kontrolle des Gens, in das das Intron integriert wurde. Daher kann man sich alle regulatorischen Elemente stromaufwärts bzw. stromabwärts der Gene oder des Gens von Interesse in dem Intron sparen. Dadurch kann man die Konstrukte entsprechend klein halten und ist sicher, dass eine Transkription auch in der betrachteten Spezies tatsächlich funktioniert. Die 30 Nutzung heterologer regulatorischer Elemente beinhaltet ein Risiko, dass diese nicht in den betrachteten Plastiden der entsprechenden Pflanzenart funktionell sind. Die Nutzung homologer Sequenzen kann aufgrund der Sequenzduplikation zu spontanen Rekombinationsereignissen mit den endogenen 35 Sequenzen und so zu einer Instabilität des Organellgenoms führen. Durch die Möglichkeit auf das Einbringen regulatorischer Elemente weitgehend verzichten zu können - beispielsweise indem man das Gen von Interesse in einem Intron kodiert, welches in eine transkriptionell aktive Region 40 des Plastoms insertiert wird, können mit dem erfindungsgemäßen Verfahren in dieser Ausführungsform zusätzlich zur Steigerung der Insertions- und Verbreitungsfähigkeit des 45

43

Transformationskonstruktes viele weitere Nachteile der konventionellen Plastidentransformation umgangen werden.

Darüber hinaus sind alle Introns nutzbar, wenn man gleichzeitig 5 die entsprechenden Faktoren, die das Spleißen vermitteln, in den Plastiden exprimiert oder sie in diese importiert. Bevorzugt sind die das Spleißen unterstützenden Faktoren im Intron selbst kodiert. Besonders bevorzugt werden in dieser Ausführungsform Introns der Gruppe II, die selbst wenigstens einen der für das 10 Spleißen notwendigen Faktoren kodieren. Dazu gehört das Ll.ltrB Intron aus *Lactococcus*. Ebenfalls bevorzugt sind Introns, die natürlicherweise in Plastiden höherer Pflanzen vorkommen, besonders Introns der Gruppe II, ganz besonders bevorzugt Introns, die für ein Protein kodieren, am meisten bevorzugt 15 Introns der *trnK* Gene des Plastidengenoms. Im letzteren Fall sind besonders bevorzugt die Introns aus den *trnK* Genen der Plastiden aus den Arten *Arabidopsis*, Mais und Tabak.

Bevorzugt sind Introns mit einer selbst-spleißenden Aktivität, 20 die nicht von weiteren Proteinfaktoren abhängt, oder Introns, die allgemeine Faktoren zum Spleißen benutzen, die universell und damit auch in Plastiden vorhanden sind, sowie Introns, die selbst für Faktoren kodieren, die für das Spleißen notwendig sind.. Zu diesen Introns gehören beispielsweise

25

- a) das Intron der Gruppe I aus *Tetrahymena* (GenBank Acc.-No.: X54512; Kruger K et al. (1982) Cell 31:147-157; Roman J und Woodson SA (1998) Proc Natl Acad Sci USA 95:2134-2139)
- 30 b) das rII-Intron der Gruppe II aus *Scenedesmus obliquus* (GenBank Acc.-No.: X17375.2 Nukleotide 28831 bis 29438; Holländer V und Kück U (1999) Nucl Acids Res 27: 2339-2344; Herdenberger F et al. (1994) Nucl Acids Res 22: 2869-2875; Kück U et al. (1990) Nucl Acids Res 18:2691-2697).
- 35 c) das Ll.LtrB Intron (GenBank Acc.-No.: U50902 Nukleotide 2854 bis 5345)
- d) das *trnK*-Intron aus *Arabidopsis* (GenBank Acc.-No.: AP000423)
- 40 e) das *trnK*-Intron aus Mais (GenBank Acc.-No.: X86563 Nukleotide komplementär 1421 bis 3909)
- f) das *trnK*-Intron aus Tabak (GenBank Acc.-No.: Z00044 Nukleotide komplementär 1752 bis 4310)

Sowohl artfremde als auch natürlicherweise in den Plastiden der jeweiligen Pflanze vorkommende Introns können genutzt werden. Zur Vermeidung von durch Sequenzduplikation bedingte Instabilitäten sind artfremde Introns - beispielsweise artfremde trnK-Introns - bevorzugt. Natürlicherweise in den Plastiden der jeweiligen Pflanze vorkommende Introns werden in einer bevorzugten Ausführungsform so modifiziert, dass sie zwar ihre Funktion noch erfüllen können, die Sequenzhomologie jedoch geringer als 95 %, bevorzugt 80 %, besonders bevorzugt 70 % zu der Sequenz des Ausgangsintrons ist.

In einer weiteren bevorzugten Ausführungsform steht ein Faktor, welcher das Spleißen des betrachteten Intron bewirkt, in trans zur Verfügung, d.h. ist nicht in dem Intron selbst kodiert. Ist dieser Faktor nicht natürlicherweise in dem betrachteten Plastid vorhanden, sondern wird erst in dieses eingebracht, kann dies auf verschiedene dem Fachmann bekannte Arten erfolgen. Beispielsweise sei das Einbringen einer entsprechenden kodierenden und exprimierbaren Sequenz in das Plastom oder das Einbringen in die nukleäre DNA genannt, wobei in letzterem Fall der Faktor vorzugsweise an eine PLS fusioniert wird.

Besonders bevorzugt sind Introns, die natürlicherweise ein DSB-Enzym (insbesondere eine Homing Endonuklease kodieren). Besonders bevorzugt ist das Intron Cp.LSU2 aus *Chlamydomonas pallidostigmatica*, welches das Enzym I-CpaI codiert (Turmel M et al. (1995) Mol Biol Evol 12:533-545). Bevorzugt sind ferner die Introns der Gruppe II aus den Mitochondrien der Hefe.

30 In einer bevorzugten Ausführungsform wird die Intronsequenz dem Insertionsort angepasst, so dass sie an diesem Lokus spleißen können. Die Anpassung kann bei Gruppe I Introns die internal guide sequence (IGS) bzw. bei den Introns der Gruppe II die exon binding sequence (EBS) I oder/und II betreffen.

35 Im Fall des trnK Introns aus Mais ist zu beachten, dass das vom trnK Intron kodierte Protein, welches auch die Maturasefunktion umfasst, in seiner natürlich kodierten Form vermutlich nicht ohne Editierung funktionell ist. Es konnte gezeigt werden, dass eine 40 Editierung (His420Tyr) der entsprechenden mRNA in Plastiden der Gerste erfolgt (Vogel J et al. (1997) J Mol Biol 270:179-187). Tyrosin an Position 420 des matK Proteins ist hoch konserviert. Auch bei den Monokotyledonen Reis und Mais wurde in der kodierenden DNA an entsprechender Position ein Kodon kodierend für His 45 gefunden. Es ist daher davon auszugehen, dass - wie in Gerste - auch in diesen Pflanzen das matK Transkript editiert wird. Da jedoch gegebenenfalls andere Pflanzenarten diese Editierung der

45.

RNA nicht gewährleisten können, wird in einer bevorzugten Ausführungsform das matK Gen im trnK-Intron aus Mais bereits durch einen entsprechenden His/Tyr-Austausch auf DNA-Ebene modifiziert, so dass eine RNA-Editierung nicht mehr erforderlich ist. Beispielsweise kann die Sequenz CATTATCATAGTGGAT des Mais trnK Intron mutiert werden zu CATTATTATAGTGGAT.

Die Spleißstelle wird im Fall von Gruppe I Introns bestimmt durch die Paarung der IGS mit dem 5' und/oder 3' zum Intron gelegenen Exon des entsprechenden Transkriptes (Lambowitz AM & Belfort M (1993) Annu Rev Biochem 62:587-622). Durch dem Fachmann bekannte Techniken wie PCR oder synthetisches Erstellen von Nukleotidsequenzen können die IGS beliebiger Gruppe I Introns entsprechend so angepasst werden, dass ein Spleißen an der vordefinierten Insertionsstelle innerhalb der DSB-Erkennungsregion erfolgt. Die veränderte IGS wird derart gestaltet, dass sie - zumindest teilweise - Basenpaarung mit den Sequenzen des Transkriptes 5' und 3' der Insertionsstelle eingehen kann. Bevorzugt wird das Intron CpLSU2 aus *C. pallidostigmatica* genutzt, welches für die Homing Endonuklease I-CpaI codiert. Wird dieses in Zusammenhang mit der Expression des DSBI-Enzyms I-CpaI genutzt, wodurch es zu einer Insertion der zu transformierenden DNA in die 23SrDNA des Plastidengenoms höherer Pflanzen kommt, ist keine Anpassung des Introns notwendig. Die Insertion erfolgt an einen Lokus im Plastidengenom höherer Pflanzen, der homolog zu dem ist, an dem sich das Intron natürlicherweise in *C. pallidostigmatica* befindet. Dieses Intron ist daher bereits derart gestaltet, dass Paarungen mit dem 5' und 3' Exon eingegangen werden können und korrektes Spleißen in dieser Nukleotidumgebung erfolgt. Bevorzugt wird ferner das Gruppe I Intron aus *Tetrahymena thermophila*. Dort unterbricht es als 413bp lange "Intervening Sequence" (IVS) die 26S rRNA codierende Region (Accession V01416 J01235 Nukleotide 53 bis 465). Die natürlicherweise zu findende IGS mit der Sequenz 5'-ggaggg-3' (Waring RB et al. 1985 Cell 40: 371-380; Been, MD & Cech, TR 1986 Cell 47: 207-216) kann durch den Fachmann bekannte Techniken an die neue Insertionsstelle angepasst werden. Wird beispielsweise eine Integration in die DSB-Erkennungsstelle des I-CpaI Enzyms an die mit ^ gekennzeichnete Stelle gewünscht (cggtcct^aaggtagcgaattc), kann die mutierte, angepasste IGS beispielsweise folgende Sequenz besitzen: 5'-gggacc-3'.

Bei Gruppe II Introns, die mobil sind, sind neben der Maturase häufig weitere Aktivitäten in dem Proteinanteil des Ribonukleoprotein-Komplexes kodiert. Diese sind für das beschriebene Verfahren jedoch nicht zwangsläufig notwendig und können daher deletiert werden. Sie werden sogar bevorzugt deletiert, da dadurch das entsprechende Konstrukt kleiner und leichter handhabbar

ist. Dem Fachmann sind verschiedene Möglichkeiten bekannt, wie die entsprechenden Aktivitäten aus dem Proteinanteil entfernt werden können. Dies kann zum Beispiel durch Erstellen eines synthetischen Gens, welches nur noch die gewünschten Bereiche 5 umfasst, oder durch geeignete PCR Methoden erfolgen.

Selbst-spleißende Introns der Gruppe II besitzen eine konser-  
vierte Struktur und bestehen im allgemeinen aus 6 verschiedenen  
Domänen. Domäne I beinhaltet die Exon-Bindestellen (EBS1 und  
10 EBS2), die beim Spleißvorgang eine Interaktion mit dem 5' vom  
Intron gelegenen Exon eingehen. Darüber hinaus findet eine Inter-  
aktion zwischen der "δ Region" (unmittelbar 5' der EBS1 gelegen)  
und der "δ' Region" am 3' Exon statt (Lambowitz AM & Belfort M  
(1993) Annu Rev Biochem 62:587-622; Michel F & Ferat JL (1995)  
15 Annu Rev Biochem 64:435-461). Diese Sequenzen können durch den  
Fachmann bekannte Techniken wie synthetisches Erstellen der  
Introns oder geeignete PCR Methoden jeweils derart angepasst  
werden, dass eine korrekte Wahl der Spleißstellen an dem in der  
DSB-Erkennungsregion gewählten Insertionsort gewährleistet ist.  
20 Dies geschieht, indem man die genannten Regionen so modifiziert,  
dass Basenpaarungen mit den entsprechenden Sequenzen stromauf-  
wärts (intron binding sequences, IBS) und stromabwärts (δ') der  
künstlichen Insertionsequenz eingegangen werden können. Wählt  
man als Insertionsort (^) für das Ll.LtrB Intron in der I-CpaI  
25 Erkennungsregion beispielsweise cggtcctaagg^agcgaaattc, so kann  
die δ und EBS1 Region beispielsweise die Sequenz TCGCTACCTTAG  
(natürliche Sequenz: TTATGGTTGTG), die EBS2 beispielsweise die  
Sequenz GACCG (natürliche Sequenz: ATGTG) annehmen. Wählt man  
das trnK Intron aus Arabidopsis thaliana, so kann die δ und EBS1  
30 Region bei Annahme der gleichen Insertionsstelle wie für das  
Ll.LtrB Intron angegeben beispielsweise die Sequenz CGCTACCTTAGG  
(natürliche Sequenz: AATGTTAAAAA) annehmen.

Wenn es sich bei der DSB-Erkennungssequenz um eine natürliche,  
35 endogene Erkennungssequenz einer Homing Endonuklease handelt,  
erfolgt die Insertion eines gewählten Introns bevorzugt an der  
Stelle der DSB-Erkennungsregion, an der auch natürlicherweise das  
Intron zugehörig zur betrachteten Homing Endonuklease zu finden  
ist.

40 Bevorzugt wird die künstliche Insertionsstelle eines Introns in  
der DSB-Erkennungsstelle derart ausgewählt, dass 5' und 3' des  
insertierten Intron möglichst viele Basen denen an der natür-  
lichen Insertionsstelle des betrachteten Introns entsprechen  
45 und die DSB-Erkennungssequenz nach der Insertion des Introns  
nicht mehr funktionell ist. Ganz besonders bevorzugt entspricht  
das jeweils unmittelbar stromaufwärts bzw. stromabwärts der

47

Insertionsstelle des Intron gelegene Nukleotid dem am natürlichen Insertionsort."

In einer besonders bevorzugten Ausführungsform ist das Intron 5 von Homologiesequenzen flankiert, um eine gerichtete Insertion zu ermöglichen. Die Homologiesequenzen sind dabei - wie oben beschrieben - homolog zu den die DSB-Erkennungssequenz flankierenden Sequenzen und ermöglichen so eine exakte Insertion.

10 Ein weiterer Gegenstand der Erfindung betrifft daher DNA-Konstrukte umfassend mindestens eine Nukleinsäuresequenz sowie Intronsequenzelemente, die befähigt sind, in einer von besagtem DNA-Konstrukt abgeleiteten Ribonukleinsäuresequenz die Deletion des für besagte Nukleinsäuresequenz kodierenden Ribonukleinsäure-15 reframentes zu gewährleisten, wobei besagte die Nukleinsäuresequenz in Bezug auf besagte Intronsequenzelemente heterolog ist.

In einer bevorzugten Ausführungsform ist die Nukleinsäuresequenz zumindesten von einer Spleißakzeptorsequenz und einer Spleißdonorsequenz flankiert.

In einer weiteren Ausführungsform umfasst das DNA-Konstrukt am 5'- und 3'-Ende Sequenzen H1 bzw. H2, die eine ausreichende Länge und Homologie zu plastidären Sequenzen H1' bzw. H2' aufweisen, um 25 eine homologe Rekombination zwischen H1 und H1' bzw. H2 und H2' und damit eine Insertion der von H1 und H2 flankierten Sequenz in das Plastom zu gewährleisten.

Ein weiterer Gegenstand der Erfindung betrifft eine transgene 30 Plastiden-DNA umfassend mindestens eine Nukleinsäuresequenz sowie Intronsequenzelemente, die befähigt sind, in einer von besagter transgenen Plastiden-DNA abgeleiteten Ribonukleinsäuresequenz die Deletion des für besagte Nukleinsäuresequenz kodierenden Ribonukleinsäurefragmentes zu gewährleisten, wobei die besagte Nukleinsäuresequenz in Bezug auf besagte Intronsequenzelemente heterolog ist. In einer bevorzugten Ausführungsform ist die Nukleinsäuresequenz zumindesten von einer Spleißakzeptorsequenz und einer Spleißdonorsequenz flankiert.

40 Die Insertionssequenz oder das Transformationskonstrukt kann zum Aufbau eines Transformationsvektor in einen Standard-Vektor wie pBluescript oder pUC18 kloniert werden. In einer bevorzugten Ausführungsform wird die Insertionssequenz bzw. Transformationskonstrukt als lineares oder linearisiertes 45 DNA-Molekül appliziert.

48

Bevorzugt wird nur der Teil des Transformationsvektors appliziert, der die Insertionssequenz oder das Transformationskonstrukt mit ggf. Homologiesequenzen, Selektionsmarker und/oder die Expressionskassette für das DSBI-Enzym umfasst.

5 Wird ganz oder teilweise auf Homologiesequenzen verzichtet, so wird das linearisierte DNA Molekül bevorzugt durch Verdau mit Restriktionsendonukleasen erhalten, die an einem oder beiden Enden einzelsträngige DNA-Überhänge generieren, die zu denen kompatibel sind, die das DSBI-Enzym in der plastidären DNA

10 erzeugt.

Der Transformationsvektor kann in einer bevorzugten Ausführungsform Elemente (z.B. einen plastidären ORI *origin of replication*, Replikationsursprung) umfassen, die es ihm ermöglichen, vor 15 der Integration in die plastidäre DNA im Plastid autonom zu replizieren oder stabil als extrachromosomal DNA-Molekül in den Plastiden zu existieren. Entsprechende Verfahren sind dem Fachmann bekannt (US 5,693,507; US 5,932,479; WO 99/10513). Dieses Verfahren ist bevorzugt, da es die zur Integration im Plastid zur 20 Verfügung stehende Kopienzahl der Insertionssequenzen steigert.

Eines der oben beschriebenen Konstrukte kann mit einem der beschriebenen Verfahren in die Plastiden einer entsprechenden Masterpflanze eingebracht werden. Bevorzugt sind Mikroinjektion 25 und besonders bevorzugt Partikelbeschuss.

#### Klonierungs-, Expressions-, Selektions- und Transformationsverfahren

30 "Expressionskassette" meint - zum Beispiel in Bezug auf die Expressionkassette für das DSBI-Enzym - solche Konstruktionen bei denen die zu exprimierende DNA in funktioneller Verknüpfung mit mindestens einem genetischen Kontrollelement steht, dass ihre Expression (d.h. Transkription und/oder Translation) ermöglicht 35 oder reguliert. Dabei kann die Expression zum Beispiel stabil oder transient, konstitutiv oder induzierbar erfolgen. Für die Einführung stehen dem Fachmann verschiedene unten aufgeführte direkte (z.B. Transfektion, Partikelbeschuss, Mikroinjektion) oder indirekte Verfahren (z.B. Agrobakterieninfektion, Virus- 40 infektion) zur Verfügung, die weiter unten aufgeführt werden.

Unter einer funktionellen Verknüpfung versteht man allgemein eine Anordnung in der eine genetische Kontrollsequenz ihre Funktion in Bezug auf eine Nukleinsäuresequenz - beispielsweise kodierend für 45 ein DSBI-Enzym - ausüben kann. Funktion kann dabei beispielsweise die Kontrolle der Expression d.h. Transkription und/oder Translation der Nukleinsäuresequenz - beispielsweise kodierend für

ein DSBI-Enzym - bedeuten. Kontrolle umfasst dabei beispielsweise die Initierung, Steigerung, Steuerung oder Suppression der Expression d.h. Transkription und ggf. Translation. Die Steuerung wiederum kann beispielsweise gewebe- und oder zeitspezifisch erfolgen. Sie kann auch induzierbar zum Beispiel durch bestimmte Chemikalien, Stress, Pathogene etc. sein.

Unter einer funktionellen Verknüpfung versteht man zum Beispiel die sequentielle Anordnung eines Promoters, der zu exprimierenden Nukleinsäuresequenz - beispielsweise kodierend für ein DSBI-Enzym - und ggf. weiterer regulativer Elemente wie zum Beispiel einem Terminator derart, dass jedes der regulativen Elemente seine Funktion bei der Expression der Nukleinsäuresequenz - beispielsweise kodierend für ein DSBI-Enzym - erfüllen kann. Dabei ist es nicht zwingend notwendig, dass die funktionelle Verknüpfung bereits auf den Transformationskonstrukten gegeben ist. Die funktionelle Verknüpfung kann sich auch in Folge der Insertion in die Kern- bzw. plastidäre DNA ergeben, wobei hier die regulativen Elemente bereits in der Kern- bzw. plastidäre DNA vorliegen. Die regulativen Elemente können diesbezüglich natürlicherweise vorliegen oder aber in einem vorgesetzten Schritt - beispielsweise bei der Einführung einer künstlichen DSB-Erkennungssequenz - eingebracht werden.

Dazu ist nicht unbedingt eine direkte Verknüpfung im chemischen Sinne erforderlich. Genetische Kontrollsequenzen, wie zum Beispiel Enhancer-Sequenzen, können ihre Funktion auch von weiter entfernten Positionen oder gar von anderen DNA-Molekülen aus auf die Zielsequenz ausüben. Bevorzugt sind Anordnungen, in denen die zu exprimierende Nukleinsäuresequenz - beispielsweise kodierend für ein DSBI-Enzym - hinter einer als Promoter fungierende Sequenz positioniert wird, so dass beide Sequenzen kovalent miteinander verbunden sind. Bevorzugt ist dabei der Abstand zwischen der Promotorsequenz und der Nukleinsäuresequenz - beispielsweise kodierend für ein DSBI-Enzym - geringer als 200 Basenpaare, besonders bevorzugt kleiner als 100 Basenpaare, ganz besonders bevorzugt kleiner als 50 Basenpaare.

Dem Fachmann sind verschiedene Wege bekannt, um zu einem der erfindungsgemäßen Transformationskonstrukte, diese umfassenden Vektoren oder einer der Expressionskassetten zu gelangen. Die Herstellung kann mittels gängiger Rekombinations- und Klonierungstechniken realisiert werden, wie sie beispielsweise in Maniatis T, Fritsch EF und Sambrook J, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in Silhavy TJ, Berman ML und Enquist LW, Experiments with Gene Fusions, Cold Spring Harbor

Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel FM et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987) beschrieben sind. Bevorzugt ist die direkte Fusion einer als Promoter fungierenden Nuklein-  
5 säuresequenz mit einer zu exprimierenden Nukleotidsequenz - beispielsweise kodierend für ein DSBI-Enzym.

Der Begriff der "genetischen Kontrollsequenzen" ist breit zu verstehen und meint all solche Sequenzen, die einen Einfluss  
10 auf das Zustandekommen oder die Funktion einer Expressionskassette oder Transformationsvektors haben. Genetische Kontrollsequenzen gewährleisten die Transkription und gegebenenfalls Translation in Zellkern (bzw. Zytoplasma) oder Plastiden. Vorzugsweise umfassen die erfindungsgemäßen Expressionskassetten  
15 5'-stromaufwärts von der jeweiligen zu exprimierenden Nuklein-säuresequenz einen Promotor und 3'-stromabwärts eine Terminatorsequenz als zusätzliche genetische Kontrollsequenz, sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils funktionell verknüpft mit der zu exprimierenden Nuklein-  
20 säuresequenz.

Genetische Kontrollsequenzen sind beispielsweise beschrieben bei "Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990)" oder "Gruber and Crosby,  
25 in: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, eds.: Glick and Thompson, Chapter 7, 89-108" sowie den dort aufgewiesenen Zitaten.

Beispiele für derartige Kontrollsequenzen sind Sequenzen, an  
30 die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Kontrollsequenzen oder anstelle dieser Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und gegebenenfalls genetisch verändert worden  
35 sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene erhöht wurde. Die Expressionskassette kann aber auch einfacher aufgebaut sein, das heißt, es werden keine zusätzlichen Regulationssignale vor die vorstehend erwähnten Gene insertiert und der natürliche Promotor mit seiner Regulation wird  
40 nicht entfernt. Stattdessen wird die natürliche Kontrollsequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert wird. Diese veränderten Promotoren können auch allein vor die natürlichen Gene zur Steigerung der Aktivität gebracht werden.

## 51

Je nach nachstehend näher beschriebenen Wirtsorganismus oder Ausgangsorganismus, der durch Einbringen der Expressionskassetten oder Vektoren in einen genetisch veränderten oder transgenen Organismus überführt wird, eignen sich verschiedene Kontrollsequenzen.

zur nukleären Expression (beispielsweise einer viralen/Bakteriophagen RNA-Polymerase oder eines DSBI-Enzyms mit einem plastidären Transitpeptid) sind grundsätzlich alle Promotoren geeignet,  
10 der die Expression von Genen, insbesondere Fremdgenen, in Pflanzen steuern können.

Geeignet sind Promotoren, die eine konstitutive Expression in Pflanzen ermöglichen (Benfey et al. (1989) EMBO J. 8:2195-2202).  
15 Vorzugsweise verwendet man insbesondere einen pflanzlichen Promotor oder einen Promotor, der einem Pflanzenvirus entstammt. Besonders bevorzugt ist der Promotor des 35S-Transkriptes des Blumenkohlmosaikvirus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985)  
20 Virology 140:281-288; Gardner et al. 1986, Plant Mol. Biol. 6, 221-228) oder den 19S CaMV Promotor (US 5,352,605 and WO 84/02913). Dieser Promotor enthält bekanntlich unterschiedliche Erkennungssequenzen für transkriptionale Effektoren, die in ihrer Gesamtheit zu einer weitgehend permanenten und konstitutiven Expression des eingeführten Gens führen (Benfey et al.  
25 (1989) EMBO J 8:2195-2202). Ein weiterer geeigneter konstitutiver Promotor ist der "Rubisco small subunit (SSU)"-Promotor (US 4,962,028). Ein weiteres Beispiel eines geeigneten Promotors ist der LeguminB-Promotor (GenBank Acc.-No.: X03677). Weitere bevorzugte konstitutive Promotoren sind zum Beispiel der Promotor der Nopalinsynthase aus Agrobakterium, der TR-Doppelpromotor, der OCS (Octopin Synthase) Promotor aus Agrobakterium, der Ubiquitin Promotor (Holtorf S et al. (1995) Plant Mol Biol 29:637-649), die Promotoren der vakuolärer ATPase Untereinheiten, der FBPaseP  
30 Promotor (WO 98/18940) oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991). Weiterhin geeignete und im Rahmen dieser Erfindung bevorzugte konstitutive Promotoren sind der SuperPromotor (Ni M et al. (1995) Plant J 7:661-676; US 5,955,646) sowie der Nitrilase-1 Promotor des ntl Gens aus  
35 Arabidopsis (GenBank Acc.-No.: Y07648.2, Nukleotide 2456 bis 4340; Hillebrand H et al. (1998) Plant Mol Biol 36 (1):89-99; Hillebrand H et al. (1996) Gene 170(2):197-200).

Bevorzugt sind induzierbare, besonders bevorzugt chemisch  
40 induzierbare Promotoren (Aoyama T und Chua NH (1997) Plant J 11:605-612; Caddick MX et al. (1998) Nat. Biotechnol 16:177-180; Review: Gatz (1997) Annu Rev Plant Physiol Plant Mol Biol

## 52

48:89-108), durch die die Expression zu einem bestimmten Zeitpunkt gesteuert werden kann. Beispielsweise seien zu nennen der PRP1 Promotor (Ward et al. (1993) Plant Mol Biol 22:361-366), ein durch Salicylsäure induzierbarer Promotor (WO 95/19443), ein durch Benzolsulfonamid-induzierbarer Promotor (EP-A-0388186), ein durch Tetrazyklin-induzierbarer Promotor (Gatz et al. (1992) Plant J 2:397-404), ein durch Abscisinsäure-induzierbarer Promotor (EP-A 335 528), ein durch Ethanol induzierbarer Promotor (Salter MG et al. (1998) Plant J. 16:127-132), der schwermetall-induzierbarer Metallothionein I Promotor (Amini S et al. (1986) Mol Cell Biol 6:2305-2316), der steroid-induzierbarer MMTV LTR Promotor (Izant JG et al. (1985) Science 229:345-352) sowie ein durch Cyclohexanon-induzierbarer Promotor (WO 93/21334). Besonders bevorzugt ist die induzierbare Expression eines PLS/DSBI-Enzym-Fusionsproteins im Kern. Induzierbare Promotoren umfasst auch solche, die durch bestimmte Repressorproteine (z.B. tet, lac) reguliert werden können. Entsprechende Repressorproteine können in Fusion mit PLS in die Plastiden translokaliert werden und dort die Expression bestimmter Gene unter Kontrolle entsprechender Promotoren regulieren. Der Repressor bindet in den Plastiden an eine künstliche ins Plastom eingefügte Repressorbindestelle und kann so die Expression des stromabwärts gelegenen Gens unterdrücken (vgl. WO95/25787). Dadurch kann beispielsweise die Expression eines plastidkodierten DSBI-Enzym im Bedarfsfall induziert oder sie bis auf den Moment, in dem die Expression gewünscht ist, unterdrückt werden.

Ferner sind Promotoren bevorzugt, die durch biotischen oder abiotischen Stress induziert werden wie beispielsweise der pathogen-induzierbare Promotor des PRP1-Gens (Ward et al., Plant Mol Biol 1993, 22: 361-366), der hitzeinduzierbare hsp70-Promotor oder der hsp80-Promoter aus Tomate (US 5,187,267), der kälteinduzierbare alpha-Amylase Promoter aus der Kartoffel (WO 96/12814) oder der verwundungsinduzierte pinII-Promoter (EP-A 0 375 091).

In einer besonders bevorzugten Ausführungsform wird vor allem die für das DSBI-Enzym kodierende Nukleinsäure unter Kontrolle eines induzierbaren Promotors exprimiert. Damit wird eine kontrollierte, steuerbare Expression erreicht und etwaige Probleme durch eine konstitutive Expression eines DSBI-Enzyms vermieden.

Vorteilhafte Kontrollsequenzen für die erfindungsgemäßen Expressionskassetten oder Vektoren umfassen virale, bakteriophage oder bakterielle Promotoren wie cos-, tac-, trp-, tet-, phoA-, tat-, lpp-, lac-, lacIq-, T7-, T5-, T3-, gal-, trc-, ara-, SP6-.

53

$\lambda$ -PR- oder im  $\lambda$ -PL-Promotor. Diese werden bevorzugt in Kombination mit der Expression der jeweils korrespondierenden RNA-Polymerase eingesetzt.

- 5 Die Expression in Plastiden kann unter Verwendung plastidärer Promotoren und/oder Transkriptionsregulationselemente realisiert werden. Beispielhaft jedoch nicht einschränkend seien zu nennen der RNA-Polymerase Promotor (WO 97/06250) oder die in WO 00/07431, US 5,877,402, WO 97/06250, WO 98/5559,5 WO 99/46394,
- 10 WO 01/42441 und WO 01/07590 beschriebenen Promotoren. Zu nennen sind das rpo B Promotorelement, das atpB Promotorelement, das clpP Promotorelement (siehe auch WO 99/46394) oder das 16S-rDNA Promotorelement. Dabei kann dem Promotor auch ein polycistronisches "Operon" zugeordnet sein (EP-A 1 076 095;
- 15 WO 00/20611). Beschrieben sind auch Systeme bei denen eine nicht-pflanzliche (z.B. virale) RNA-Polymerase unter Verwendung plastidärer Transitpeptide in das Plastid importiert wird und dort spezifisch die Expression transgener Sequenzen induziert, die unter Kontrolle der Erkennungssequenzen der RNA-Polymerase
- 20 stehen und zuvor in die plastidäre DNA insertiert wurden (WO 95/16783; US 5,925,806; US 5,575,198).

Neben den genannten Promotoren wären weiter bevorzugt zu verwenden:

- 25 a) der PrbcL Promotor (SEQ ID NO: 44)
- b) der Prps16 Promotor (SEQ ID NO: 50)
- 30 c) der Prrn16 Promotor (SEQ ID NO: 46)

In einer besonders bevorzugten Ausführungsform werden NEP Promotoren eingesetzt. Dies sind Promotoren, die in Plastiden funktionell sind und von der nuklear kodierten, plastidären RNA-Polymerasen (NEP) erkannt werden. Bevorzugt sind: Prrn-62, Pyrf2-1577; PatpB-289; Prps2-152; Prps16-107; Pyrf1-41; PatpI-207; PclpP-511; PclpP-173 und PaccD-129 (WO 97/06250; Hajdukiewicz PTJ et al (1997) EMBO J 16:4041-4048).

40 Besonders bevorzugt sind:

- a) Der PaccD-129 Promotor des accD Gens aus Tabak (WO 97/06250; SEQ ID NO: 47)
- 45 b) Der PclpP-53 Promotor des clpP Gens als hoch-aktiver NEP Promotor in Chloroplasten (WO 97/06250; SEQ ID NO: 48)

- c) Der Prrn-62 Promotor des rrn Gens (SEQ ID NO: 49)
- d) Der Prps16-107 Promotor des rps16 Gens (SEQ ID NO: 45):
- 5 e) Der PatpB/E-290 Promotor des atpB/E Gens aus Tabak (Kapoor S et al. (1997) Plant J 11:327-337) (SEQ ID NO: 51)
- f) Der PrpoB-345 Promotor des rpoB Gens (Liere K & Maliga P (1999) EMBO J 18: 249-257) (SEQ ID NO: 52)

10

Im Allgemeinen sind in dieser bevorzugten Ausführungsform alle Promotoren nutzbar, die der Klasse III angehören (Hajdukiewicz PTJ et al (1997) EMBO J 16:4041-4048) sowie alle Fragmente der Promotoren der Klasse II, die die Transkriptionsinitiation durch 15 die NEP steuern. Solche Promotoren bzw. Promotor-Anteile sind nicht besonders hoch konserviert. Als Consensus nahe der Transkriptionsinitiationsstelle von NEP Promotoren wird angegeben: ATAGAATAAA (Hajdukiewicz PTJ et al (1997) EMBO J 16:4041-4048).

20 Normalerweise werden Gene von regulatorischen Sequenzen umgeben, die aus den Plastiden des zu transformierenden Organismus stammen. Dadurch erzeugt man Sequenzduplikationen, die zu Instabilitäten aufgrund von spontanen, intrachromosomalen homologen Rekombinationseignissen führen können (Heifetz PB (2000) Biochimie 82(6-7):655-666). Es wurde zur Lösung dieses Problems vorgeschlagen, heterologe regulatorische Sequenzen zu nutzen, oder 25 endogen im Plastidengenom bereits vorhandene regulatorische Einheiten auszunutzen (WO 99/46394; WO 01/42441). Auch eine Verminderung der Homologie durch Mutagenese der endogenen Promotorsequenz ist beschrieben (WO 01/07590).

Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen wie die oben genannten für das erfindungsgemäße Verfahren verwendet werden. Besonders bevorzugt sind Promotoren isoliert aus Prokaryoten. Ganz besonders bevorzugt sind 35 Promotoren isoliert aus Synechocystis oder E.coli. Darüber hinaus können auch synthetische Promotoren vorteilhaft verwendet werden, wie beispielsweise ein synthetischer Promotor abgeleitet von der E. coli Konsensus Sequenz für σ<sub>70</sub> Promotoren:

40

5'-TTGACA N<sub>16-19</sub> TATAAT N<sub>3</sub> CAT -3',

wobei N für jedes beliebige Nukleotid (also A, G, C oder T) steht. Dem Fachmann ist offensichtlich, dass auch einzelne bis 45 wenige Basenaustausche in den angegebenen, konservierten Regionen möglich sind, ohne die Funktion des Promotors zu zerstören. Die variable Gestaltung dieser synthetischen Promotoren durch Nutzen

## 55

verschiedener Sequenzabfolgen, ermöglicht es, eine Vielzahl von Promotoren zu erstellen, die nicht extensive Homologien aufweisen, was insbesondere für den Fall, dass mehrere Promotoren benötigt werden, die Stabilität der Expressionskassetten im Plastom erhöht. Beispielsweise sind besonders bevorzugte Promotorsequenzen genannt, die von oben genannter Konsensussequenz abgeleitet sind:

- a) 5'-TTGACATTCACTCTTCAATTATCTATAATGATACA-3' (SEQ ID NO: 53)  
10 b) 5'-TTGACAATTTCCTCTGAATTATATAATTAAACAT-3' (SEQ ID NO: 72)

Dem Fachmann ist offensichtlich, daß diese synthetischen Promotoren die Expression beliebiger Gene steuern können. Sie können beispielsweise dazu genutzt werden, die Expression eines Selektionsmarkers anzutreiben – auch, um unter regenerativen Bedingungen auf transplastome Pflanzen unter Zuhilfenahme von besagtem Selektionssystem selektieren zu können. Selektionsmarker sind weiter unten beispielhaft aufgezählt. Darüber hinaus können solche synthetischen Promotoren mit jedem beliebigen Gen verknüpft werden, beispielsweise mit Genen kodierend für Antikörper, Antigene oder Enzyme. Bevorzugt enthalten die Expressionskassetten bestehend aus solchen Promotoren auch im folgenden näher beschriebene 5'-untranslatierte Regionen (oder Ribosomenbindestellen) oder 3'-nichtkodierende Regionen.

25

Ein weiterer Gegenstand der Erfindung betrifft Expressionskassetten enthaltend eine für ein DSBI-Enzym kodierende Nukleinsäuresequenz unter Kontrolle eines in pflanzlichen Plastiden funktionellen Promotors, beispielsweise eines der oben beschriebenen Promotoren. Die Expressionskassette kann weitere Elemente, wie beispielsweise Transkriptionsterminatoren und/oder Selektionsmarker beinhalten.

Genetische Kontrollsequenzen umfassen auch weitere Promotoren, 35 Promotorelemente oder Minimalpromotoren, die die expressionsteuernden Eigenschaften modifizieren können. Genetische Kontrollsequenzen umfassen ferner auch die 5'-untranslatierte Region (5'-UTR) oder die nichtkodierende 3'-Region (3'-UTR) von Genen (Eibl C (1999) Plant J 19: 1-13). Es ist gezeigt worden, dass 40 diese eine signifikante Funktionen bei der Regulation der Genexpression in Plastiden höherer Pflanzen spielen können. Auch im Kern können genetische Kontrollelemente wie 5'-UTR, Introns oder 3'-UTR eine Funktion bei der Genexpression besitzen. So wurde beispielweise gezeigt, dass 5'-untranslatierte Sequenzen die trans-45 iente Expression heterologer Gene verstärken können. Sie können

56

ferner die Gewebespezifität fördern (Rouster J et al., Plant J., 1998, 15: 435-440.).

Bevorzugt werden als 5' UTRs und 3'UTRs in Plastiden eingesetzt:

5

- a) 5'psbA (aus Tabak) (SEQ ID NO: 54)
- b) 5'rbcL einschließlich 5' Anteilen aus dem codierenden Bereich des rbcL Gens (aus Tabak) (SEQ ID NO: 55); die mit SEQ ID NO: 55 beschriebene Sequenz wurden gegenüber der nativen Sequenz mutiert, um eine PstI bzw. NcoI Schnittstelle einzubringen.
- c) 5'rbcLs (SEQ ID NO: 56); die mit SEQ ID NO: 56 beschriebene Sequenz wurden gegenüber der nativen Sequenz mutiert, um eine PstI Schnittstelle einzubringen.
- d) 3'psbA-1 aus Synechocystis (SEQ ID NO: 57)
- 20 e) 3'psbA aus Tabak (SEQ ID NO: 58)
- f) 3'rbcL aus Tabak (SEQ ID NO: 59)

Genetische Kontrollsequenzen vor allem für die Expression in Plastiden umfassen insbesondere auch Ribosomenbindungssequenzen zur Initiation der Translation. Diese sind für gewöhnlich in den 5'UTRs enthalten. Dies ist vor allem dann bevorzugt, wenn von der zu exprimierenden Nukleinsäuresequenz entsprechende Sequenzen nicht bereitgestellt werden oder diese mit dem Expressions-  
30 system kompatibel sind. Insbesondere bevorzugt ist die Verwendung einer synthetischen Ribosomenbindestelle (RBS) mit der Sequenz 5'-GGAGG(N)<sub>3-10</sub>ATG-3', bevorzugt 5'-GGAGG(N)<sub>5</sub>ATG-3' (SEQ ID NO: 60) besonders bevorzugt 5'-GGAGGATCTCATG-3' (SEQ ID NO: 61).

35

Die Expressionskassette kann vorteilhafterweise eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promoter enthalten, die eine erhöhte transgene Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der transgen zu exprimierenden Nukleinsäuresequenzen können zusätzliche vorteilhafte Sequenzen insertiert werden, wie weitere regulatorische Elemente oder Terminatoren. Die transgen zu exprimierenden Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

45

57.

Ferner ist es möglich, nach dem Startcodon eine sogenannte Downstream-Box einzufügen, die die Expression im Allgemeinen steigert (Translations- "Enhancer" WO 00/07431; WO 01/21782).

- 5 Als genetische Kontrollsequenzen v.a. bei der Kerntransformation geeignete Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA Polyadenylierungssignale aus *Agrobakterium tumefaciens*, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids  
10 pTiACHS entsprechen (Gielen et al., EMBO J. 3 (1984), 835 ff) oder funktionelle Äquivalente davon. Beispiele für besonders geeignete Terminatorsequenzen sind der OCS (Octopin-Synthase)-Terminator und der NOS (Nopaline-Synthase)-Terminator.
- 15 Die erfindungsgemäßen Transformationsvektoren und Insertionssequenzen können weitere Nukleinsäuresequenzen umfassen. Solche Nukleinsäuresequenzen können bevorzugt Expressionskassetten darstellen. Beispielhaft aber nicht einschränkend für die in den Expressionskonstrukten zu exprimierenden DNA-Sequenzen  
20 seien zu nennen:

1. Selektionsmarker

"Selektionsmarker" meint all solche Nukleinsäure- oder Proteinsequenzen, deren Expression (d.h. Transkription und ggf. Translation) einer Zelle, Gewebe oder Organismus einen anderen Phänotyp verleiht, als einer nicht transformierten. Selektionsmarker umfasst beispielsweise solche Nukleinsäure- oder Proteinsequenzen, deren Expression einer Zelle, Gewebe oder Organismus  
30 einen Vorteil (positiver Selektionsmarker) oder Nachteil (negativer Selektionsmarker) gegenüber Zellen vermittelt, die diese Nukleinsäure oder Protein nicht exprimieren. Positive Selektionsmarker wirken beispielsweise dadurch, dass sie auf die Zelle inhibitorisch wirkende Substanz detoxifiziert wird (Bsp.  
35 Antibiotika-/Herbizidresistenz), oder eine Substanz gebildet wird, welche der Pflanze unter den gewählten Bedingungen verbesserte Regeneration oder erhöhtes Wachstum ermöglicht (zum Beispiel nutritive Marker, hormonproduzierender Marker wie ipt, s.u.). Eine andere Form positiver Selektionsmarker umfasst  
40 mutierte Proteine oder RNAs, die gegenüber einem selektiven Agenten nicht empfindlich sind (beispielsweise 16S rRNA Mutanten, die unempfindlich gegenüber Spectinomycin sind). Negative Selektionsmarker wirken beispielsweise dadurch, dass sie die Bildung einer toxischen Substanz in den transformierten Zellen katalysieren  
45 (zum Beispiel das codA Gen). Ferner kann Selektionsmarker auch Reporterproteine umfassen, insofern diese geeignet sind, transformierte von nicht-transformierten Zellen, Geweben oder Organen

(beispielsweise durch Farbgebung oder einen anderen detektierbaren Phänotyp) zu unterscheiden.

Nachfolgende Selektionsmarker seien beispielhaft und nicht einschränkend zu nennen:

1.1 Positive Selektionsmarker:

Der mit der Expressionskassette in den Zellkern oder die 10 Plastiden eingebrachte selektionierbare Marker verleiht den erfolgreich transformierten Zellen eine Resistenz gegen ein Biozid (zum Beispiel ein Herbizid wie Phosphinothricin, Glyphosat oder Bromoxynil), einen Metabolismusinhibitor wie 2-Desoxyglucose-6-phosphat (WO 98/45456) oder ein Antibiotikum, wie zum 15 Beispiel Tetracycline, Ampicillin, Kanamycin, G'418, Neomycin, Bleomycin oder Hygromycin. Der Selektionsmarker erlaubt die Selektion der transformierten Zellen von untransformierten (McCormick et al., Plant Cell Reports 5 (1986), 81-84; Dix PJ & Kavanagh TA (1995) Euphytica 85: 29-34).

20

Besonders bevorzugte Selektionsmarker sind solche, die eine Resistenz gegen Herbicide verleihen. Beispielhaft als Selektionsmarker seien genannt:

- 25 - DNA Sequenzen, die für Phosphinothricinacetyltransferasen (PAT) kodieren, welche die freie Aminogruppe des Glutamin synthaseinhibitors Phosphinothricin (PPT) acetylieren und damit eine Detoxifizierung des PPT erreichen (de Block et al. 1987, EMBO J. 6, 2513-2518) (auch Bialaphos® resistenzgen (bar) genannt). Das bar Gen kodierend für eine Phosphinothricinacetyltransferase (PAT) kann aus beispielsweise Streptomyces hygroscopicus oder S. viridochromogenes isoliert werden. Entsprechende Sequenzen sind dem Fachmann bekannt (aus Streptomyces hygroscopicus GenBank Acc.-No.: X17220 und X05822, aus Streptomyces viridochromogenes GenBank Acc.-No.: M 22827 und X65195; US 5,489,520). Ferner sind synthetische Gene beispielsweise für die Expression in Plastiden beschrieben. Ein synthetisches Pat Gen ist beschrieben in Becker et al. (1994) The Plant J. 5:299-307. 30 Die Gene verleihen Resistenz gegen das Herbizid Bialaphos oder Glufosinat und sind vielbenutzer Marker in transgenen Pflanzen (Vickers JE et al. (1996). Plant Mol Biol Reporter 14:363-368; Thompson CJ et al. (1987) EMBO J 6:2519-2523).  
35  
40  
45 - 5-Enolpyruvylshikimat-3-phosphatsynthasegene (EPSP Synthase-gene), die eine Resistenz gegen Glyphosat (N-(phosphonomethyl)glycin) verleihen. Das unselektive Herbizid Glyphosat

## 59

- hat die 5-Enolpyruyl-3-phosphoshikimatsynthase (EPSPS) als molekulares Target. Diese hat eine Schlüsselfunktion in der Biosynthese aromatischer Aminosäuren in Mikroben und Pflanzen, jedoch nicht in Säugern (Steinrucken HC et al. 5 (1980) Biochem. Biophys. Res. Commun. 94:1207-1212; Levin JG und. Sprinson DB (1964) J. Biol. Chem. 239: 1142-1150; Cole DJ (1985) Mode of action of glyphosate a literature analysis. p. 48-74. In: Grossbard E und Atkinson D (eds.). The herbicide glyphosate. Buttersworths, Boston.). Glyphosat-tolerante EPSPS Varianten werden bevorzugt als Selektionsmarker verwendet (Padgett SR et al. (1996). New weed control opportunities: development of soybeans with a Roundup Ready™ gene. In: Herbicide Resistant Crops (Duke, S.O., ed.), pp. 53-84. CRC Press, Boca Raton, FL; Saroha MK und Malik VS (1998) J Plant Biochemistry and Biotechnology 7:65-72). Das EPSPS Gen des Agrobakterium sp. strain CP4 hat eine natürliche Toleranz gegen Glyphosat, die auf entsprechende transgene Pflanzen transferiert werden kann. Das CP4 EPSPS Gen wurde aus Agrobakterium sp. strain CP4 kloniert (Padgett 10 SR et al. (1995) Crop Science 35(5):1451-1461). Sequenzen von 5-Enolpyrvylshikimate-3-phosphate-synthasen, die Glyphosat-tolerant sind, wie beispielsweise beschrieben in US 5,510,471; US 5,776,760; US 5,864,425; US 5,633,435; US 5,627,061; US 5,463,175; EP 0 218 571, sind sowohl in den Patenten beschriebenen als auch in der GenBank hinterlegt. Weitere Sequenzen sind beschrieben unter GenBank Accession X63374. Ferner ist das aroA Gen bevorzugt (M10947 S. typhimurium aroA locus 5-enolpyrvylshikimate-3-phosphate synthase (aroA protein) gene).
- 15 30. - das für das Glyphosat® degradierende Enzyme kodierende gox Gen (Glyphosatoxidoreduktase). GOX (beispielsweise die Glyphosatoxidoreductase aus Achromobacter sp.) katalysiert die Spaltung einer C-N Bindung im Glyphosat, welches so zu Aminomethylphosphonsäure (AMPA) und Glyoxylat umgesetzt wird. GOX kann dadurch eine Resistenz gegen Glyphosat vermitteln. (Padgett SR et al. (1996) J Nutr. 1996 Mar;126(3):702-16; Shah D et al. (1986) Science 233: 478-481).
- 20 35. - das deh Gen (kodierend für eine Dehalogenase, die Dalapon® inaktiviert), (GenBank Acc.-No.: AX022822, AX022820 sowie WO99/27116)
- 25 40. - bxn Gene, die für Bromoxynil degradierende Nitrilase-enzyme kodieren. Beispielsweise die Nitrilase aus Klebsiella ozanenae. Sequenzen sind in der Genbank beispielsweise unter den Acc.-No: E01313 (DNA encoding bromoxynil specific

## 60

- nitrilase) und J03196 (*K. pneumoniae* bromoxynil-specific nitrilase (bxn) gene, complete cds) zu finden.
- Neomycinphosphotransferasen verleihen eine Resistenz gegen Antibiotika (Aminoglykoside) wie Neomycin, G418, Hygromycin, Paromomycin oder Kanamycin, indem sie durch eine Phosphorylierungsreaktion deren inhibierende Wirkung reduzieren. Besonders bevorzugt ist das nptII Gen. Sequenzen können aus der GenBank erhalten werden (AF080390 Mini-transposon mTn5-GNm; AF080389 Minitransposon mTn5-Nm, complete sequence). Zudem ist das Gen bereits Bestandteil zahlreicher Expressionsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden (AF234316 pCAMBIA-2301; AF234315 pCAMBIA-2300, AF234314 pCAMBIA-2201). Das NPTII Gen kodiert für eine Aminoglycosid-3'0-phosphotransferase aus *E.coli*, Tn5 (GenBank Acc.-No: U00004 Position 1401-2300; Beck et al. (1982) Gene 19 327-336). Darüber hinaus kann auch das aphA-6 Gen aus *Acinetobacter baumannii*, kodierend für eine Aminoglykosidphosphotransferase, als Selektionsmarker genutzt werden (Huang et al. (2002) Mol Genet Genomics 268:19-27)
- das DOGR<sup>1</sup>-Gen. Das Gen DOGR<sup>1</sup> wurde aus der Hefe *Saccharomyces cerevisiae* isoliert (EP 0 807 836). Es codiert für eine 2-Desoxyglukose-6-phosphat Phosphatase, die Resistenz gegenüber 2-DOG verleiht (Randez-Gil et al. 1995, Yeast 11, 1233-1240; Sanz et al. (1994) Yeast 10:1195- 1202, Sequenz: GenBank Acc.-No.: NC001140 chromosom VIII, *Saccharomyces cervisiae* Position 194799-194056).
- Sulfonylurea- und Imidazolinon inaktivierende Acetolactat-synthasen, die eine Resistenz gegen Imidazolinon/Sulfonylurea-Herbizide verleihen. Beispielhaft seien für Imidazolinon-Herbizide die Wirkstoffe Imazamethabenz-methyl, Imazamox, Imazapyr, Imazaquin, Imazethapyr zu nennen. Für Sulfonylharnstoff-Herbizide seien beispielhaft Amidosulforon, Azimsulfuron, Chlorimuronethyl, Chlorsulfuron, Cinosulfuron, Imazosulfuron, Oxasulfuron, Prosulfuron, Rimsulfuron, Sulfo-sulfuron zu nennen. Dem Fachmann sind zahlreiche weitere Wirkstoffe der genannten Klassen bekannt. Geeignet sind Nukleinsäuresequenzen wie beispielsweise die unter der Gen-Bank Acc-No.: X51514 abgelegte Sequenz für das *Arabidopsis thaliana* Csr 1.2 Gen (EC 4.1.3.18) (Sathasivan K et al. (1990) Nucleic Acids Res. 18(8):2188). Acetolactatesynthasen,

## 61

die eine Resistenz gegen Imidazolinon-Herbizide verleihen,  
sind ferner beschrieben unter den GenBank Acc.-No.:

- 5 a) AB049823 Oryza sativa ALS mRNA for acetolactate synthase,  
complete cds, herbicide resistant biotype
- b) AF094326 Bassia scoparia herbicide resistant acetolactate  
synthase precursor (ALS) gene, complete cds
- 10 c) X07645 Tobacco acetolactate synthase gene, ALS SuRB  
(EC 4.1.3.18)
- d) X07644 Tobacco acetolactate synthase gene, ALS SuRA  
(EC 4.1.3.18)
- 15 e) A19547 Synthetic nucleotide mutant acetolactate synthase
- f) A19546 Synthetic nucleotide mutant acetolactate synthase
- 20 g) A19545 Synthetic nucleotide mutant acetolactate synthase
- h) I05376 Sequence 5 from Patent EP 0257993
- i) I05373 Sequence 2 from Patent EP 0257993
- 25 j) AL133315

Hygromycinphosphotransferasen (X74325 P. pseudomallei gene  
for hygromycin phosphotransferase) die eine Resistenz gegen  
das Antibiotikum Hygromycin verleihen. Das Gen ist Bestand-  
teil zahlreicher Expressionsvektoren und kann unter Verwen-  
dung von dem Fachmann geläufigen Verfahren (wie beispiels-  
weise Polymerasekettenreaktion) aus diesen isoliert werden  
(AF294981 pINDEX4; AF234301 pCAMBIA-1380; AF234300 pCAM-  
BIA-1304; AF234299 pCAMBIA-1303; AF234298 pCAMBIA-1302;  
35 AF354046 pCAMBIA-1305.; AF354045 pCAMBIA-1305.1)

- Resistenzgene gegen
- 40 a) Chloramphenicol (Chloramphenicolacetyltransferase),
  - b) Tetracyclin, verschiedene Resistenzgene sind beschrieben  
z.B. X65876 S. ordonez genes class D tetA and tetR  
for tetracycline resistance and repressor proteins  
X51366 Bacillus cereus plasmid pBC16 tetracycline  
resistance gene. Zudem ist das Gen bereits Bestandteil  
zahlreicher Expressionsvektoren und kann unter Verwendung

62

von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden

- 5           c) Streptomycin, verschiedene Resistenzgene sind beschrieben z.B. mit der GenBank Acc.-No.: AJ278607 *Corynebacterium acetoacidophilum* ant gene for streptomycin adenylyl-transferase.
- 10          d) Zeocin, das entsprechende Resistenzgen ist Bestandteil zahlreicher Klonierungsvektoren (z.B. L36849 Cloning vector pZEO) und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden.
- 15          e) Ampicillin ( $\beta$ -Lactamase Gen; Datta N, Richmond MH. (1966) Biochem J. 98(1):204-9; Hefron F et al. (1975) J. Bacteriol 122: 250-256; das Amp Gen wurde zuerst zur Herstellung des *E. coli* Vektors pBR322 kloniert; Bolivar F et al. (1977) Gene 2:95-114). Die Sequenz ist Bestandteil zahlreicher Klonierungsvektoren und kann unter Verwendung von dem Fachmann geläufigen Verfahren (wie beispielsweise Polymerasekettenreaktion) aus diesen isoliert werden.
- 20          25 - Gene wie die Isopentenyltransferase aus Agrobakterium tumefaciens (strain: P022) (Genbank Acc.-No.: AB025109). Das ipt Gen ist ein Schlüsselenzym der Cytokinin-Biosynthese. Seine Überexpression erleichtert die Regeneration von Pflanzen (z.B. Selektion auf Cytokinin-freiem Medium). Das Verfahren zur Nutzung des ipt Gens ist beschrieben (Ebinuma H et al. (2000) Proc Natl Acad Sci USA 94:2117-2121; Ebinuma, H et al. (2000) Selection of Marker-free transgenic plants using the oncogenes (ipt, rol A, B, C) of Agrobakterium as selectable markers, In Molecular Biology of Woody Plants. 35 Kluwer Academic Publishers).

Verschiedene weitere positive Selektionsmarker, die den transformierten Pflanzen einen Wachstumsvorteil gegenüber nicht-transformierten verleihen, sowie Verfahren zu ihrer Verwendung sind 40 u.a. beschrieben in EP-A 0 601 092. Beispielhaft sind zu nennen  $\beta$ -Glucuronidase (in Verbindung mit z.B. Cytokininglucuronid), Mannose-6-phosphat-Isomerase (in Verbindung mit Mannose), UDP-Galaktose-4-Epimerase (in Verbindung mit z.B. Galactose), wobei Mannose-6-phosphat-Isomerase in Verbindung mit Mannose besonders 45 bevorzugt ist.

Für einen in Plastiden funktionellen Selektionsmarker sind insbesondere solche bevorzugt, die eine Resistenz gegen Spectinomycin, Streptomycin, Kanamycin, Lincomycin, Gentamycin, Hygromycin, Methotrexat, Bleomycin, Phleomycin, Blasticidin, Sulfonamid-, Phosphinotricin, Chlorsulfuron, Bromoxymil, Glyphosat, 2,4-Datrazin, 4-methyltryptophan, Nitrat, S-aminoethyl-L-cysteine, Lysin/Threonin, Aminoethyl-Cystein oder Betainaldehyd verleihen. Besonders bevorzugt sind die Gene *aadA*, *nptII*, *BADH*, *FLARE-S* (eine Fusion aus *aadA* und GFP, beschrieben bei Khan MS & Maliga 10 P, 1999 Nature Biotech 17: 910-915).

Als in Plastiden funktionelle Selektionsmarker ist vor allem das *aadA* Gen beschrieben (Svab Z und Maliga P (1993) Proc Natl Acad Sci USA 90:913-917). Ferner beschrieben sind modifizierte 16S 15 rDNA, das *nptII* Gen (Kanamycinresistenz) und das *bar* Gen (Phosphinotricinresistenz). Aufgrund der Bevorzugung des Selektionsmarker *aadA* wird dieser bevorzugt "recycled" d.h. nach seiner Verwendung aus dem Genom bzw. Plastom deletiert (Fischer N et al. (1996) Mol Gen Genet 251:373-380; Corneille S et al. (2001) Plant 20 J 27:171-178), so dass *aadA* in weiteren Transformationen einer bereist transplastomen Pflanze wieder als Selektionsmarker benutzt werden kann. Als weiterer möglicher Selektionsmarker ist die Betaine-aldehyd-Dehydrogenase (*BADH*) aus Spinat beschrieben (Daniell H et al. (2001) Trends Plant Science 6:237-239; Daniell 25 H et al. (2001) Curr Genet 39:109-116; WO 01/64023; WO 01/64024; WO 01/64850). Auch lethal wirkende Agenzen wie beispielsweise Glyphosat können in Verbindung mit entsprechend detoxifizierenden oder resistenten Enzymen genutzt werden (WO 01/81605).

30 Es können auch *binding type marker* genutzt werden. Die Nutzen der als Insertionsstelle bevorzugten DBS-Erkennungssequenz der Homing-Endonuklease I-CpaI im Gen der 23S rRNA, wird wenigstens das 3'-Ende der Insertionssequenz (bevorzugt eines artifiziellen Introns) mit homologen Sequenzen der Zielregion umgeben. Es 35 werden also Sequenzen der 23SrDNA in den Transformationsvektor aufgenommen. An einer Stelle (Position 2073 oder 2074 der 23SrRNA aus Tabak, Sequenz: AAAGACCCTATGAAG) können Punktmutationen eingeführt werden (z.B. Sequenz: GGAGACCCTATGAAG), die den aus einer so mutierten 23SrDNA abgeleiteten Ribosomen Resistenz 40 gegenüber Lincomycin vermitteln (Cseplö A et al. (1988) Mol Gen Genet 214:295-299). Weitere Punktmutation umfassen solche in der 16S rRNA von Tabak, die Resistenz gegenüber Spectinomycin verleihen (Mutation unterstrichen):

45 a) 5'-GGAAGGTCAGGATGC-3' ("nativ" hier: A)

Andere Mutationen verleihen eine Resistenz gegen Streptomycin:

- b) 5'-GAATGAAACTA-3' ("nativ" hier: C)

### 5 1.2 Negative Selektionsmarker

Negative Selektionsmarker ermöglichen beispielsweise die Selektion von Organismen mit erfolgreich deletierten Sequenzen, die das Markergen umfassen (Koprek T et al. (1999) The Plant Journal 10 19(6):719-726). Beispielsweise können so für Selektionsmarker oder für DSBI-Enzyme kodierende Sequenzen nach erfolgreicher Anwendung des erfindungsgemäßen Verfahrens wieder aus dem Genom/ Plastom deletiert werden.

- 15 Bei der negativen Selektion wird beispielsweise durch den in die Pflanze eingebrachten negativen Selektionsmarker eine Verbindung, die ansonsten für die Pflanze keine nachteilige Wirkung hat, in eine Verbindung mit nachteiliger Wirkung umgesetzt. Ferner sind Gene geeignet, die per se eine nachteilige Wirkung haben, wie zum Beispiel TK thymidine kinase (TK) und Diphtheria Toxin A Fragment (DT-A), das codA Genprodukt kodierend für eine Cytosindeaminase (Gleave AP et al. (1999) Plant Mol Biol 40(2):223-35; Perera RJ et al. (1993) Plant Mol Biol 23(4): 793-799; Stougaard J (1993) Plant J 3:755-761), das Cytochrom P450 Gen (Koprek et al. (1999) 20 Plant J. 16:719-726), Gene kodierend für eine Haloalkan Dehalogenase (Naested H. (1999) Plant J. 18:571-576), das iaaH Gen (Sundaresan V et al. (1995) Genes & Development 9:1797-1810) oder das tms2 Gen (Fedoroff NV & Smith DL (1993) Plant J 3:273-289).
- 25 25 Plant J. 16:719-726), Gene kodierend für eine Haloalkan Dehalogenase (Naested H. (1999) Plant J. 18:571-576), das iaaH Gen (Sundaresan V et al. (1995) Genes & Development 9:1797-1810) oder das tms2 Gen (Fedoroff NV & Smith DL (1993) Plant J 3:273-289).
- 30 30 Die jeweils für die Selektion verwendeten Konzentrationen der Antibiotika, Herbizide, Biozide oder Toxine müssen an die jeweiligen Testbedingungen bzw. Organismen angepasst werden. Beispielhaft seien für Pflanzen zu nennen Kanamycin (Km) 50 bis 100 mg/L, Hygromycin B 40 mg/L, Phosphinothricin (Ppt) 6 bis 20 35 mg/L, Sepctinomycin (Spec) 15 bis 500 mg/L.

Ferner können funktionelle Analoga der genannten Nukleinsäuren kodierend für Selektionsmarker exprimiert werden. Funktionelle Analoga meint hier all die Sequenzen, die im wesentlichen die gleiche Funktion haben d.h. zu einer Selektion transformierter 40 Organismen befähigt sind. Dabei kann das funktionelle Analogon sich in anderen Merkmalen durchaus unterscheiden. Es kann zum Beispiel eine höhere oder niedrigere Aktivität haben oder auch über weitere Funktionalitäten verfügen.

Funktionelle Analoga meint ferner Sequenzen, die für Fusionsproteine kodieren bestehend aus einem der bevorzugten Selektionsmarker und einem anderen Protein zum Beispiel einem weiteren bevorzugten Selektionsmarker, einem der unten genannten Reporterproteine oder einer PLS. Beispielhaft sei eine Fusion des GFP (grün fluoreszierenden Proteins) und des aadA Gens genannt (Sidorov VA et al. (1999) Plant J 19:209-216).

## 2. Reportergene

Reportergene kodieren für leicht quantifizierbare Proteine, die über Eigenfarbe oder Enzymaktivität eine Bewertung der Transformationseffizienz, des Expressionsortes oder -zeitpunktes oder die Identifizierung transgener Pflanzen gewährleisten. Ganz besonders bevorzugt sind dabei Gene kodierend für Reporter-Proteine (siehe auch Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) wie

- "green fluorescence protein" (GFP) (Chui WL et al., Curr Biol 1996, 6:325-330; Leffel SM et al., Biotechniques. 23(5):912-8, 1997; Sheen et al. (1995) Plant Journal 8(5):777-784; Haseloff et al. (1997) Proc Natl Acad Sci USA 94(6):2122-2127; Reichel et al. (1996) Proc Natl Acad Sci USA 93(12):5888-5893; Tian et al. (1997) Plant Cell Rep 16:267-271; WO 97/41228).
- Chloramphenicoltransferase,
- Luziferase (Millar et al., Plant Mol Biol Rep 1992 10:324-414; Ow et al. (1986) Science, 234:856-859); erlaubt Biolumineszenzdetektion.
- $\beta$ -Galactosidase, kodiert für ein Enzym für das verschiedenen chromogene Substrate zur Verfügung stehen.
- $\beta$ -Glucuronidase (GUS) (Jefferson et al., EMBO J. 1987, 6, 3901-3907) oder das uidA Gen, das ein Enzym für verschiedene chromogene Substrate kodiert.
- R-Locus Genprodukt:Protein, das die Produktion von Anthocyaninpigmenten (rote Färbung) in pflanzlichen Gewebe reguliert und so eine direkte Analyse der Promoteraktivität ohne Zugabe zusätzlicher Hilfsstoffe oder chromogener Substrate ermöglicht (DellaPorta et al., In: Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium, 11:263-282, 1988).

66

- $\beta$ -Lactamase (Sutcliffe (1978) Proc Natl Acad Sci USA 75:3737-3741), Enzym für verschiedene chromogene Substrate (z.B. PADAC, eine chromogenes Cephalosporin).
  - 5 - xylE Genprodukt (Zukowsky et al. (1983) Proc Natl Acad Sci USA 80:1101-1105), Catecholdioxygenase, die chromogene Catechole umsetzen kann.
  - Alpha-Amylase (Ikuta et al. (1990) Bio/technol. 8:241-242).
  - 10 - Tyrosinase (Katz et al. (1983) J Gen Microbiol 129:2703-2714), Enzym, das Tyrosin zu DOPA und Dopaquinon oxidiert, die infolge das leicht nachweisbare Melanin bilden.
  - 15 - Aequorin (Prasher et al. (1985) Biochem Biophys Res Commun 126(3):1259-1268), kann in der Calcium-sensitiven Biolumineszenzdetektion verwendet werden.
- Der Selektionsmarker bzw. das Reportergen ist bevorzugt auf dem Transformationskonstrukt, besonders bevorzugt auf der Insertionssequenz kodiert. Er kann aber auch auf einem unabhängigen Transformationskonstrukt kodiert sein, welches in einer Co-Transformation mit dem Transformationskonstrukt von Interesse in den Kern oder die Plastiden einer Pflanzenzelle eingebracht wird.
- 25 Das erfindungsgemäße Transformationsvektoren und Insertionssequenzen können weitere Funktionselemente enthalten. Der Begriff der weitere Funktionselemente ist breit zu verstehen. Bevorzugt sind all solche Elemente gemeint, die einen Einfluss auf Herstellung, Vermehrung, Funktion, Nutzen oder Wert der im Rahmen des erfindungsgemäßen Verfahrens zum Einsatz kommenden Insertionssequenzen, Transformationskonstrukte oder -vektoren haben. Beispielhaft jedoch nicht einschränkend seien für die weiteren Funktionselemente zu nennen:
- 35 i. Replikationsursprünge (ORI; origin of DNA replication), die eine Vermehrung der erfindungsgemäßen Expressionskassetten oder Vektoren in zum Beispiel E. coli oder aber auch in Plastiden gewährleisten. Beispielhaft für E.coli ORIs seien genannt der pBR322 ori, der P15A ori (Sambrook et al.: Molecular Cloning. A Laboratory Manual, 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) oder der colE1-ORI beispielsweise aus pBLUESCRIPT. Plastidäre ORI sind beschrieben in US 5,693,507, US 5,932,479 oder WO 99/10513.
- 40
- 45

67.

- ii. Multiple Klonierungsregionen (MCS) erlauben und erleichtern die Insertion eines oder mehrerer Nukleinsäuresequenzen.
- iii. Sequenzen, die eine homologe Rekombination bzw. Insertion in das Genom bzw. Plastom eines Wirtsorganismus ermöglichen.
- iv. Elemente zum Beispiel "Bordersequenzen", die einen Agrobakterien-vermittelte Transfer in Pflanzenzellen für die Übertragung und Integration ins Pflanzengenom ermöglichen, wie zum Beispiel die rechte oder linke Begrenzung der T-DNA oder die vir-Region.

Die Einführung einer Insertionssequenz oder eines Expressionskonstruktes für ein DSBI-Enzym kann vorteilhaft unter Verwendung von Vektoren realisiert werden, in die diese Konstrukte bzw.

Kassetten insertiert werden. Vektoren können beispielhaft Plasmide, Cosmide, Phagen, Viren, Retroviren oder auch Agrobacterien sein.

20 In einer vorteilhaften Ausführungsform wird die Einführung der Expressionskassette mittels Plasmidvektoren realisiert. Bevorzugt sind solche Vektoren, die eine stabile Integration der Expressionskassette in das Wirtsgenom bzw. -plastom ermöglichen.

25 Die Herstellung eines transformierten Organismus oder einer transformierten Zelle erfordert, dass die entsprechende DNA in die entsprechende Wirtszelle bzw. in die Plastiden derselben eingebracht wird. Für diesen Vorgang, der als Transformation bezeichnet wird, steht eine Vielzahl von Methoden zur Verfügung (siehe auch Keown et al. (1990) Methods in Enzymology 185:527-537). So kann die DNA beispielhaft direkt durch Mikroinjektion, Elektroporation oder durch Bombardierung mit DNA-be- schichteten Mikropartikeln eingeführt werden. Auch kann die Zelle chemisch, zum Beispiel mit Polyethylenglycol, permeabilisiert

30 werden, so dass die DNA durch Diffusion in die Zelle gelangen kann. Die Transformation kann auch durch Fusion mit anderen DNA- enthaltenden Einheiten wie Minicells, Zellen, Lysosomen oder Liposomen erfolgen. Zu nennen sind ferner die Transfektion unter Einsatz von Calciumphosphat, DEAE-Dextran oder kationischen

40 Lipiden, Transduktion, Infektion, die Inkubation trockener Embryonen in DNA-haltiger Lösung, die Sonikation sowie die Transformation intakter Zellen oder Gewebe durch Mikro- oder Makro- injektion in Gewebe oder Embryonen, Gewebeelektroporation oder die Vakuumfiltration von Samen. Derartige Verfahren sind dem

45 Fachmann geläufig. Im Falle von Injektion oder Elektroporation von DNA in pflanzliche Zellen sind keine besonderen Anforderungen an das verwendete Plasmid gestellt. Einfache Plasmide wie die der

pUC-Reihe können verwendet werden. Sollen vollständige Pflanzen aus den transformierten Zellen regeneriert werden, so ist es nützlich, dass sich auf dem Plasmid ein zusätzliches selektionierbares Markergen befindet. Auch die Verfahren zur Regeneration von Pflanzen aus Pflanzengeweben oder Pflanzenzellen sind beschrieben.

Es gibt verschiedene Möglichkeiten, die DNA in die Plastiden einzuschleusen. Für die vorliegende Erfindung ist lediglich entscheidend, dass die DNA in die Plastiden eingebracht wird. Die vorliegende Erfindung ist jedoch nicht limitiert auf ein bestimmtes Verfahren. Jedes Verfahren, das ein Einbringen der zu transformierenden DNA in die Plastiden einer höheren Pflanze erlaubt, ist geeignet. Die stabile Transformation von Plastiden ist ein dem Fachmann geläufiges Verfahren und wurde für höhere Pflanzen beschrieben (u.a. bei Svab Z und Maliga P (1993) Proc Natl Acad Sci USA 90(3):913-917). Die Methoden beruhen zum Beispiel auf einer Transformation mittels "Particle Gun" und einer Insertion in das plastidäre Genom durch homologe Rekombination unter Selektionsdruck. Weitere Methoden sind beschrieben in US 5,877,402. In EP-A 0 251 654 wird die DNA mittels Agrobakterium tumefaciens eingeführt (s. De Block M et al. (1985) EMBO J 4:1367-1372; Venkateswarlu K und Nazar RN (1991) Bio/Technology 9:1103-1105). Ferner wurde gezeigt, dass man mittels Elektroporation DNA in isolierte Chloroplasten einbringen und so eine transiente Expression erzielen kann (To KY et al. (1996) Plant J 10:737-743). Bevorzugt ist eine Transformation mittels eines direkten Transfers der DNA in Plastiden von Protoplasten beispielsweise unter Verwendung von PEG (Polyethylenglykol) (Koop HU et al. (1996) Planta 199:193-201; Kofer W et al. (1998) In Vitro Cell Dev Biol Plant 34:303-309; Dix PJ und Kavanagh TA (1995) Euphytica. 85:29-34; EP-A 0 223 247). Am meisten bevorzugt sind biolistische Transformationsverfahren. Dabei wird die zu transformierende DNA auf z.B. Gold- oder Wolframpartikel aufgebracht. Diese Partikel werden anschließend auf das zu transformierende Explantat beschleunigt (Dix PJ und Kavanagh TA (1995) Euphytica. 85:29-34; EP-A 0 223 247). Anschließend werden in der dem Fachmann geläufigen Weise transplastome Pflanzen unter Selektionsdruck auf geeignetem Medium regeneriert. Entsprechende Verfahren sind beschrieben (z.B. US 5,451,513; US 5,877,402; Svab Z et al. (1990) Proc Natl Acad Sci USA 87:8526-8530; Svab Z und Maliga P (1993) Proc Natl Acad Sci USA 90:913-917). Darüber hinaus kann die DNA mittels Mikroinjektion in die Plastiden eingebracht werden. Ein besonderes Verfahren der Mikroinjektion wurde kürzlich beschrieben (Knoblauch M et al. (1999) Nature Biotech 17:906-909; van Bel AJE et al. (2001) Curr Opin Biotechnol 12:144-149). Dieses Verfahren ist für die vorliegende Erfindung

69.

besonders bevorzugt. Es ist auch möglich, durch Protoplastenfusion die Plastiden aus einer Art in eine andere Art einzubringen, diese dort zu transformieren und anschließend durch Protoplastenfusion diese wieder in die ursprüngliche Art zu 5 überführen (WO 01/70939).

Neben diesen "direkten" Transformationstechniken kann eine Transformation auch durch bakterielle Infektion mittels Agrobakterium tumefaciens oder Agrobakterium rhizogenes durchgeführt werden

- 10 (Horsch RB (1986) Proc Natl Acad Sci USA 83(8):2571-2575; Fraley et al. (1983) Proc Natl Acad Sci USA 80:4803-4807; Bevans et al. (1983) Nature 304:184-187). Die Expressionskassette beispielsweise für das DSBI-Enzym wird bevorzugt in spezielle Plasmide integriert, entweder in einen Zwischenvektor (englisch: shuttle 15 or intermediate vector) oder einen binären Vektor. Bevorzugt werden binäre Vektoren verwendet. Binäre Vektoren können sowohl in *E. coli* als auch in Agrobakterium replizieren und können direkt in Agrobakterium transformiert werden (Holsters et al. (1978) Mol Gen Genet 163:181-187). Verschiedene binäre Vektoren 20 sind bekannt und teilweise kommerziell erhältlich wie zum Beispiel pBIN19 (Clontech Laboratories, Inc. USA; Bevan et al. (1984) Nucl Acids Res 12:8711). Das Selektionsmarkergen erlaubt eine Selektion transformierter Agrobakterien und ist zum Beispiel das nptII Gen, das eine Resistenz gegen Kanamycin verleiht. 25 Das Binärplasmid kann beispielsweise durch Elektroporation oder andere Transformationsmethoden in den Agrobakterienstamm übertragen werden (Mozo & Hooykaas 1991, Plant Mol. Biol. 16, 917-918). Die Cokultur der pflanzlichen Explantate mit dem Agrobakterienstamm findet in der Regel für zwei bis drei Tage statt. 30 Das in diesem Fall als Wirtsorganismus fungierende Agrobakterium sollte bereits ein Plasmid mit der vir-Region enthalten. Viele Stämme von Agrobakterium tumefaciens sind in der Lage, genetisches Material zu übertragen, wie z.B. die Stämme EHA101[pEHA101] (Hood EE et al. (1996) J Bacteriol 35 168(3):1291-1301), EHA105[pEHA105] (Hood et al. (1993) Transgenic Research 2:208-218), LBA4404[pAL4404] (Hoekema et al. (1983) Nature 303:179-181), C58C1[pMP90] (Koncz and Schell (1986) Mol Gen Genet 204:383-396) und C58C1[pGV2260] (Deblaere et al. (1985) Nucl Acids Res 13: 4777-4788).

- 40 Für den Transfer der DNA auf die pflanzliche Zelle werden pflanzliche Explantate mit Agrobakterium tumefaciens oder Agrobakterium rhizogenes kokultiviert. Ausgehend von infiziertem Pflanzenmaterial (z.B. Blatt-, Wurzel- oder Stengelteile, aber 45 auch Protoplasten oder Suspensionen von Pflanzenzellen) können ganze Pflanzen unter Verwendung eines geeigneten Mediums, dass zum Beispiel Antibiotika oder Biozide zur Selektion trans-

formierten Zellen enthalten kann, regeneriert werden. Ein mit-transformierter Selektionsmarker erlaubt die Selektion von transformierten Zellen von untransformierten (McCormick et al. (1986) Plant Cell Reports 5:81-84). Die erhaltenen Pflanzen können in üblicher Weise gezüchtet, geselbstet und gekreuzt werden. Zwei oder mehr Generationen sollten kultiviert werden, um sicherzustellen, dass die genomische Integration stabil und vererblich ist. Die genannten Verfahren sind beispielsweise beschrieben bei Jenes B et al. (1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von Kung SD und Wu R, Academic Press, S.128-143 sowie bei Potrykus (1991) Ann Rev Plant Physiol Plant Molec Biol 42:205-225).

15 Die Agrobakterium-vermittelte Transformation ist am besten für dicotyledone Pflanzenzellen geeignet, wohingegen die direkten Transformationstechniken sich für jeden Zelltyp eignen.

Die Agrobakterium-vermittelte Transformation wird besonders bevorzugt für die Kerntransformation, die direkten Transformationstechniken besonders bevorzugt für die Plastidentransformation eingesetzt.

Sobald eine überwiegend homotransplastome Pflanzenzelle nach dem erfindungsgemäßen Verfahren hergestellt wurde, kann eine vollständige Pflanze unter Verwendung von dem Fachmann bekannten Verfahren erhalten werden. Hierbei geht man beispielhaft von Kalluskulturen aus. Aus diesen noch undifferenzierten Zellmassen kann die Bildung von Spross und Wurzel in bekannter Weise induziert werden. Die erhaltenen Sprösslinge können ausgepflanzt und gezüchtet werden.

#### Deletionsverfahren

35 Bei den beschriebenen erfindungsgemäßen Verfahren ist es auf verschiedenen Stufen vorteilhaft, bestimmte zuvor eingeführte Sequenzen (beispielsweise für Selektionsmarker und/oder DSBI-Enzyme) wieder aus dem Plastom oder Genom der Pflanze oder Zelle zu entfernen. So ist es von Vorteil aber nicht zwingend erforderlich, den Selektionsmarker, der beispielsweise bei der Insertion einer nicht-natürlichen DBS-Erkennungssequenz eingeführt wurde, aus der Masterpflanze wieder zu entfernen, weil dann in einer nachfolgenden Transformation (z.B. mit der Insertionssequenz) wieder der gleiche Selektionsmarker genutzt werden kann. Eine 45 Deletion ist insbesondere vorteilhaft, da der Selektionsmarker nach der Selektionsphase nicht mehr unbedingt erforderlich und daher überflüssig ist. Die Deletion erhöht zudem die Verbraucher-

71

akzeptanz und ist unter zulassungstechnischen Überlegungen wünschenswert. Zudem wird der Proteinsyntheseapparat des Plastides nicht unnötig durch die Synthese des Markerproteins belastet, was sich potentiell vorteilhaft auf die Eigenschaften 5 der entsprechenden Pflanze auswirkt.

Zur gezielten Deletion von Sequenzen sind dem Fachmann verschiedenen Verfahren bekannt. Zu nennen ist beispielsweise jedoch nicht einschränkend die Excision mittels Rekombinassen.

- 10 Verschiedene sequenzspezifische Rekombinationssystemen sind beschrieben wie das Cre/lox-System des Bacteriophagen P1 (Dale EC und Ow DW (1991) Proc Natl Acad Sci USA 88:10558-10562; Russell SH et al. (1992) Mol Gen Genet 234: 49-59; Osborne BI et al. (1995) Plant J. 7, 687-701), das FLP/FRT System der Hefe (Kilby 15 NJ et al. (1995) Plant J 8:637-652; Lyznik LA et al. (1996) Nucleic Acids Res 24:3784-3789), die Gin Rekombinase des Mu Phagen, die Pin Rekombinase aus E. coli oder das R/RS System des pSR1 Plasmids (Onouchi H et al. (1995) Mol Gen Genet 247:653-660; Sugita K et al. (2000) Plant J 22:461-469). Diese Verfahren sind 20 nicht nur für die Deletion von DNA Sequenzen aus dem Kerngenom, sondern auch aus dem Plastom nutzbar (Corneille et al. (2001) Plant J 27: 171-178; Hajdukiewicz et al. (2001) Plant J 27:161-170). Als weitere Rekombinassen können beispielsweise eingesetzt werden: Phic31 (Kuhstoss & Rao (1991) J Mol Biol 25 222:897-908), TP901 (Christiansen et al. (1996) J Bacteriol 178:5164-5173), xisF aus Anabaena (Ramaswamy et al. (1997) Mol Microbiol 23:1241-1249), Integrase vom Phagen PhiLC3 (Lillehaug et al. (1997) Gene 188:129-136) oder die Rekombinase kodiert vom sre Gen des R4-Phagen (Matsuura et al. (1996) J Bacteriol 30 178:3374-3376).

- In einer bevorzugten Ausführungsform wird die Deletion jedoch durch intrachromosomalen Rekombination aufgrund von entsprechend eingebrachten Sequenzduplikationen realisiert. Letzteres kann 35 durch das gezielte Einführen von Doppelstrangbrüchen nahe der Sequenz-Duplikationen in seiner Effizienz gesteigert werden (vgl. Fig. 8). Dazu wird die zu deletierende Sequenz beidseitig von Homologiesequenzen H1 und H2 flankiert, die eine ausreichende Länge und Homologie aufweisen, um miteinander zu rekombinieren. 40 Die Rekombination wird durch die Induktion mindestens eines sequenzspezifischen Doppelstrangbruches nahe einer der beiden Homologiesequenzen gelegenen weiteren DSB-Erkennungssequenz (die jedoch bevorzugt von der ersten verschieden ist) induziert. Bevorzugt ist diese DSB-Erkennungssequenz zwischen den beiden 45 Homologiesequenzen lokalisiert. Zur Induktion des Doppelstrangbruches wird bevorzugt ein zweites DSBI-Enzym exprimiert oder eingebracht, das von dem ersten unterschieden ist. Besonders

72.

bevorzugt wird dieses Verfahren zur Deletion von Selektionsmarkern aus dem Plastom genutzt.

- Ein anderer Gegenstand der Erfindung betrifft die mit dem erfindungsgemäßen Verfahren hergestellten transplastomen, überwiegend homoplastomen Pflanzen, sowie Teile derselben wie Blätter, Wurzeln, Samen, Früchte, Knollen, Pollen oder Zellkulturen, Kallus usw. - abgeleitet von solchen.
- 10 Ein anderer Gegenstand der Erfindung betrifft die in dem erfindungsgemäßen Verfahren zum Einsatz kommenden Pflanzen, die eine erfindungsgemäße Expressionskassette für ein DSBI-Enzym oder ein Fusionsprotein aus PLS und DSBI-Enzym enthalten. Dabei liegt - besonders bevorzugt - die Expressionskassette für das Fusionsprotein aus PLS und DSBI-Enzym unter Kontrolle eines im pflanzlichen Zellkern funktionellen Promotors stabil integriert in die nukleäre DNA vor. Bevorzugt liegt die Expressionskassette kodierend für ein DSBI-Enzym unter Kontrolle eines in pflanzlichen Plastiden aktiven Promotors stabil integriert in das Plastom vor.
- 15 20 Umfasst sind ferner Teile derselben wie Blätter, Wurzeln, Samen, Knollen, Früchte, Pollen oder Zellkulturen, Kallus usw. - abgeleitet von vorgenannten Pflanzen.

Von Menschen und Tieren verzehrbare erfindungsgemäße, genetisch veränderte Pflanzen können auch beispielsweise direkt oder nach an sich bekannter Aufbereitung als Nahrungsmittel oder Futtermittel verwendet werden.

Ein weiterer Gegenstand der Erfindung betrifft die Verwendung der oben beschriebenen erfindungsgemäßen transplastomen, überwiegend homoplastomen Pflanzen und der von ihnen abgeleitete Zellen, Zellkulturen, Teile - wie zum Beispiel bei transgenen pflanzlichen Organismen Wurzeln, Blätter etc.-, und transgenes Vermehrungsgut wie Samen oder Früchte, zur Herstellung von Nahrungsmitteln, Pharmazeutika oder Feinchemikalien.

Feinchemikalien meint Enzyme wie beispielsweise die unten genannten industriellen Enzyme, Vitamine wie beispielsweise Tocopherolen und Tocotrienolen (z.B. Vitamin E) und Vitamin B2, Aminosäuren wie beispielsweise Methionin, Lysin oder Glutamat, Kohlenhydrate wie beispielsweise Stärke, Amylose, Amylopektin oder Saccharose, Fettsäuren wie beispielsweise gesättigte, ungesättigte und polyungesättigte Fettsäuren, natürliche und synthetische Geschmacks-, Aromastoffe wie beispielsweise Linalool, Menthol, Borneon (Kampfer), Pinen, Limonen oder Geraniol und Farbstoffe wie beispielsweise Retinoide (z.B. Vitamin A), Flavonoide (z.B. Quercetin, Rutin, Tangeretin, Nobiletin) oder

73

Carotinoide (z.B.  $\beta$ -Carotin, Lycopin, Astaxanthin). Besonders bevorzugt ist die Produktion von Tocopherolen und Tocotrienolen sowie Carotinoiden. Die Züchtung der transformierten Wirtsorganismen sowie die Isolierung aus den Wirtsorganismen bzw. aus dem Züchtungsmedium erfolgt mit dem Fachmann bekannten Verfahren. Die Produktion von Pharmazeutika, wie zum Beispiel Antikörpern oder Vakzinen ist beschrieben (Hood EE, Jilka JM. (1999) Curr Opin Biotechnol. 10(4):382-386; Ma JK und Vine ND (1999) Curr Top Microbiol Immunol. 236:275-92).

10

Insbesondere eignet sich das erfindungsgemäße Verfahren zur Herstellung von industriellen Enzymen im Rahmen eines sogenannten "Phytofarming". Beispielhaft jedoch nicht einschränkend seien für die industriellen Enzyme zu nennen Lipasen, Esterasen, Proteasen, Nitrilasen, Acylasen, Epoxyhydrolasen, Amidasen, Phosphatasen, Xylanasen, Alkoholdehydrogenasen, Amylasen, Glucosidasen, Galactosidasen, Pullulanasen, Endocellulasen, Glucanasen, Cellulasen, Nukleasen, Chitindeacetylasen, Monoaminoxidases, Lysozyme und Laccasen.

15

Im Rahmen der Erfindung insbesondere bevorzugte Ausführungsformen werden unten im Rahmen der Erläuterungen zu den Abbildungen näher beschrieben.

#### 25 Sequenzen

##### 1. SEQ ID NO: 1

pCB42-94 Basisvektor für die Plastidentransformation.

##### 30 2. SEQ ID NO:2

In die Multiple Klonierungsstelle von pCB42-94 (SEQ ID NO: 1) insertierte Nukleinsäuresequenz. Resultierender Vektor:  
pCB199-3.

##### 35 3. SEQ ID NO:3

In die Multiple Klonierungsstelle von pCB42-94 (SEQ ID NO: 1) insertierte Nukleinsäuresequenz. Resultierender Vektor:  
pCB401-20

##### 40 4. SEQ ID NO:4

Expressionkassette aus pCB289-13 zur plastidären Expression der I-PpoI Homing-Endonuklease.

##### 5. SEQ ID NO:5

45 Aminosäuresequenz der durch die Expressionkassette aus pCB289-13 kodierten I-PpoI Homing-Endonuklease.

6. SEQ ID NO:6  
Zur Erstellung des Vektors pCB304-25 eingesetztes XbaI / BglII - Fragment.
- 5 7. SEQ ID NO:7  
In die Multiple Klonierungsstelle von pGEMTeasy insertierte Nukleinsäuresequenz. Resultierender Vektor: pCB220-17
8. SEQ ID NO:8  
10 In die Multiple Klonierungsstelle von pBluescript insertierte Nukleinsäuresequenz. Resultierender Vektor: pCB270-1
9. SEQ ID NO:9  
15 Sequenz aus Vektor pCB315-1: LacZ-Gen mit insertiertem Intron zum Nachweis des Spleißens.
10. SEQ ID NO: 10  
L1.LtrB Intron aus Vektor pCB345-34.
- 20 11. SEQ ID NO: 11  
Synthetische Sequenz der Homing-Endonuklease I-PpoI (ORF: 16 bis 507)
12. SEQ ID NO: 12  
25 Proteinsequenz der Homing-Endonuklease I-PpoI
13. SEQ ID NO: 13  
Nukleinsäuresequenz der Homing-Endonuklease I-CpaI aus Chlamydomonas pallidostigmatica (Veränderungen zur veröffentlichten Sequenz an Pos. 69. Am ATG wurde eine NcoI Schnittstelle eingeführt. (ORF: 4 bis 462)
14. SEQ ID NO: 14  
35 Proteinsequenz der Homing-Endonuklease I-CpaI
15. SEQ ID NO: 15  
Sequenz enthaltend das CplSU2 Intron
16. SEQ ID NO: 16: Oligonukleotid-Primer p19  
40 5'-TAAGGCCCTCGGTAGCAACGG-3'
17. SEQ ID NO: 17: Oligonukleotid-Primer p20  
5'-GGGGTACCAAATCCAACTAG-3'
- 45 18. SEQ ID NO: 18: Oligonukleotid-Primer p21:  
5'-GGAGCTCGCTCCCCCGCCGTCGTTTC-3'

75

19. SEQ ID NO: 19: Oligonukleotid-Primer p22  
5'-GATGCATGATGACTTGACGGCATCCTC-3'
20. SEQ ID NO: 20: Oligonukleotid-Primer p190  
5'-GTCGACAGATCTTTAA-3'
21. SEQ ID NO: 21: Oligonukleotid-Primer p191  
5'-AGATCTGTCGACTTAA-3'
- 10 22. SEQ ID NO: 22: Oligonukleotid-Primer p199  
5'-GATCTCCAGTTAACCTGGGTAC-3'
23. SEQ ID NO: 23: Oligonukleotid-Primer p200  
5'-CCCAGTTAACTGGGA-3'
- 15 24. SEQ ID NO: 24: Oligonukleotid-Primer p218  
5'-TTAAGCCAGTTAACTGGCGGAGCT-3'
25. SEQ ID NO: 25: Oligonukleotid-Primer p219  
20 5'-CCGCCAGTTAACTGGC-3'
26. SEQ ID NO: 26: Oligonukleotid-Primer p276  
5'-TCGAGAAGATCAGCCTGTTATCCCTAGAGTAAC-3'
- 25 27. SEQ ID NO: 27: Oligonukleotid-Primer p277  
5'-CTAGAGTTACTCTAGGGATAACAGGCTGATCTTC-3'
28. SEQ ID NO: 28: Oligonukleotid-Primer p91  
5'-AGAAGACGATCCTAAGG-3'
- 30 29. SEQ ID NO: 29: Oligonukleotid-Primer p92  
5'-TGAAGACTTGACAAGGAATTTCGC-3'
- 30 30. SEQ ID NO: 30: Oligonukleotid-Primer p102  
35 5'-AGAAGACGATCCTAAATAGCAATATTACCTTGGGACCAAAAGTTATCAGGCATG-3'
31. SEQ ID NO: 31: Oligonukleotid-Primer p103  
5' TGAAGACTTGACAAGGAATTTCGCTACCTTCGAGTACTCCAAAACTAATC-3'
- 40 32. SEQ ID NO: 32: Oligonukleotid-Prime p207  
5'-GAGAAGACATTCCCTAACACACATCCATAACGTGCG-3'
33. SEQ ID NO: 33: Oligonukleotid-Prime p208  
5'-TGAAGACTTGACATTGATATGGTGAAGTAGG-3'

45

34. SEQ ID NO: 34  
Nukleinsäuresequenz kodierend für das Transitpeptid der kleinen Untereinheit (SSU) der Ribulosebisphosphatcarboxylase (Rubisco ssu) aus Erbse
- 5 35. SEQ ID NO: 35  
Transitpeptid der kleinen Untereinheit (SSU) der Ribulosebisphosphatcarboxylase (Rubisco ssu) aus Erbse
- 10 36. SEQ ID NO: 36  
Transitpeptid der plastidären Transketolase aus Tabak.
- 15 37. SEQ ID NO: 37  
Nukleinsäuresequenz kodierend für das Transitpeptid der plastidären Transketolase aus Tabak (Leseraster 1; pTP09)
- 20 38. SEQ ID NO: 38  
Nukleinsäuresequenz kodierend für das Transitpeptid der plastidären Transketolase aus Tabak (Leseraster 2; pTP10)
- 25 39. SEQ ID NO: 39  
Nukleinsäuresequenz kodierend für das Transitpeptid der plastidären Transketolase aus Tabak (Leseraster 3; pTP11)
40. SEQ ID NO: 40  
Transitpeptid der plastidären Isopentenylpyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana*.
41. SEQ ID NO: 41  
30 Nukleinsäuresequenz kodierend für das Transitpeptid der plastidären Isopentenylpyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* (Leseraster 1; IPP-9)
42. SEQ ID NO: 42  
35 Nukleinsäuresequenz kodierend für das Transitpeptid der plastidären Isopentenylpyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* (Leseraster 2; IPP-10)
43. SEQ ID NO: 43  
40 Nukleinsäuresequenz kodierend für das Transitpeptid der plastidären Isopentenylpyrophosphat Isomerase-2 (IPP-2) aus *Arabidopsis thaliana* (Leseraster 3; IPP-11)
44. SEQ ID NO: 44  
45 Nukleinsäuresequenz kodierend für den Prbcl Promotor aus Tabak.

45. SEQ ID NO: 45  
Nukleinsäuresequenz kodierend für den Prps16-107 Promotor aus Tabak.
- 5 46. SEQ ID NO: 46  
Nukleinsäuresequenz kodierend für den Prrn16 Promotor aus Tabak.
- 10 47. SEQ ID NO: 47  
Nukleinsäuresequenz kodierend für den PaccD-129 Promotor aus Tabak.
- 15 48. SEQ ID NO: 48  
Nukleinsäuresequenz kodierend für den PclpP-53 Promotor aus Tabak.
- 20 49. SEQ ID NO: 49  
Nukleinsäuresequenz kodierend für den Prrn-62 Promotor aus Tabak.
- 25 50. SEQ ID NO: 50  
Nukleinsäuresequenz kodierend für den Prps16 Promotor aus Tabak.
- 30 51. SEQ ID NO: 51  
Nukleinsäuresequenz kodierend für den PatpB/E-290 Promotor aus Tabak.
- 35 52. SEQ ID NO: 52  
Nukleinsäuresequenz kodierend für den PrpoB-345 Promotor aus Tabak.
53. SEQ ID NO: 53  
Nukleinsäuresequenz kodierend für einen Promotor abgeleitet von der Konsensussequenz der 670 Promotoren aus E.coli.
- 35 54. SEQ ID NO: 54  
Nukleinsäuresequenz kodierend für die 5'-untranslatierte Region des psbA Gens aus Tabak (5'psbA).
- 40 55. SEQ ID NO: 55  
Nukleinsäuresequenz kodierend für die 5'-untranslatierte Region einschließlich 5' Anteilen aus dem kodierenden Bereich des rbcL Gens aus Tabak (5'rbcL).

56. SEQ ID NO: 56  
Nukleinsäuresequenz kodierend für die 5'-untranslatierte Region des rbcLs Gens aus Tabak.
- 5 57. SEQ ID NO: 57  
Nukleinsäuresequenz kodierend für die 3'-untranslatierte Region des psbA-1 Gens aus Synechocystis (3'psbA-1)
58. SEQ ID NO: 58  
10 Nukleinsäuresequenz kodierend für die 3'-untranslatierte Region des psbA Gens aus Tabak (3'psbA)
59. SEQ ID NO: 59  
15 Nukleinsäuresequenz kodierend für die 3'-untranslatierte Region des rbcL Gens aus Tabak (3'rbcL)
60. SEQ ID NO: 60  
Nukleinsäuresequenz kodierend für synthetische Ribosomen-bindestellen (RBS)  
20
61. SEQ ID NO: 61  
Nukleinsäuresequenz kodierend für synthetische Ribosomen-bindestelle (RBS)
- 25 62. SEQ ID NO: 62  
Vollständiges Insert des Vektors pCB304-25
63. SEQ ID NO: 63  
BglII / MunI Fragment des Vektors pCB320-192.  
30
64. SEQ ID NO: 64: Oligonukleotid-Primer p93  
5'-AAAGATCTCCTCACAAAGGGGGTCG-3'
65. SEQ ID NO: 65: Oligonukleotid-Primer p97  
35 5'-TCGAAGACTTAGGACCGTTATAG-3'
66. SEQ ID NO: 66: Oligonukleotid-Primer p98  
5'-AGGAAGACCTTGTGGGTAAGTTCCG-3'
- 40 67. SEQ ID NO: 67: Oligonukleotid-Primer p95:  
5'-CTCAATTGGGGTCTCTGTCCAGGTGCAGG-3'
68. SEQ ID NO: 68: Nukleinsäuresequenz kodierend für Fusionsproteine aus der nativen I-Ppo-I Nuklease und der der  
45 IPP-Plastidenlokalisationssequenz (ORF für I-PpoI: 181-672; IPP-Transitpeptid: 1-180; native Sequenz von 1-172).

79

69. SEQ ID NO: 69: Fusionsproteine aus der nativen I-Ppo-I Nuklease und der der IPP-Plastidenlokalisationssequenz.
70. SEQ ID NO: 70: Nukleinsäuresequenz kodierend für lange Version der I-PpoI Homing-Endonuklease.
71. SEQ ID NO: 71: Aminosäuresequenz kodierend für lange Version der I-PpoI Homing-Endonuklease.
- 10 72. SEQ ID NO: 72: Nukleinsäuresequenz kodierend für eine Promotorsequenz abgeleitet von der Konsensussequenz der σ70 Promotoren aus E.coli.
- 15 73. SEQ ID NO: 73: Nukleinsäuresequenz kodierend für das künstliche Intron TetIVS2a.
74. SEQ ID NO: 74: Insert des Vektors pCB459-1
- 20 75. SEQ ID NO: 75: Insert des Vektors pCB478-3
76. SEQ ID NO: 76: Insert des Vektors pCB492-25
77. SEQ ID NO: 77: Oligonukleotidprimer p396  
25 5'-TAGTAAATGACAATTTCCTCTGAATTATATAATTAAACATGGCGACTGTACCAAAA  
AC-3
78. SEQ ID NO: 78: Oligonukleotidprimer p95  
30 5'-CTCAATTGGGTCTCTGTCCAGGTGCAGG-3'
79. SEQ ID NO: 79: Nukleinsäuresequenz kodierend für PCR Produkt Prom-TetIVS2a-Cpa
80. SEQ ID NO: 80: Insert des Vektors pCB435-45
- 35 81. SEQ ID NO: 81: Nukleinsäuresequenz kodierend für Sonde für Southernblot-Analyse (gerichtet gegen Teile der 16SrDNA).
- 40 82. SEQ ID NO: 82: Nukleinsäuresequenz kodierend für Sonde für Southernblot-Analyse (gerichtet gegen Teile der 23SrDNA).
83. SEQ ID NO: 83: Insert des Vektors pCB456-2.
- 45 84. SEQ ID NO: 84: Insert des Vektors pCB528-2, von KpnI bis SacI

## Abbildungen

Im Rahmen des erfindungsgemäßen Verfahrens sind insbesondere die in nachfolgenden Abbildungen verdeutlichten Ausführungsformen besonders bevorzugt. Für die Abbildungen gelten allgemein nachfolgende Abkürzungen:

- A, A' Paar homologer Sequenzen A und A'
- A/A' Ergebnis einer homologen Rekombination zwischen A und A'
- 10 und/oder eines durch Reperatursynthese bedingten Austausches zwischen A und A'.
- B, B' Paar homologer Sequenzen B und B'
- B/B' Ergebnis einer homologen Rekombination zwischen B und B'
- 15 und/oder eines durch Reperatursynthese bedingten Austausches zwischen B und B'.
- H1, H2: Paar homologer Sequenzen H1 und H2
- H1/2: Sequenz als Ergebnis der homologen Rekombination aus H1 und H2
- DS funktionelle DSB-Erkennungssequenz
- 20 nDS nicht funktionelle "Halbseite" einer DSB-Erkennungssequenz
- E: DSBI-Enzym
- P: Promotor
- I: weitere Nukleinsäuresequenz (Gen von Interesse)
- 25 S, S' Positive Selektionsmarker
- NS Negativer Selektionsmarker
- IS Intronsequenzen. Das Intron ist insgesamt als Box gekennzeichnet. Die Box umfasst alle für ein funktionelles Intron erforderlichen Elemente.
- 30 Wie bereits oben beschrieben stellt A/A' bzw. B/B' das Ergebnis einer homologen Rekombination und/oder eines durch Reperatursynthese bedingten Austausches dar. Die resultierende Sequenz kann wiederum Ausgangssequenz für weitere homologe Rekombinationen bzw. Reperatursynthesen sein. Zur Vereinfachung wird diese Sequenz (A/A' bzw. B/B') in nachfolgenden Schritten wieder als A bzw. B bezeichnet.
- 1. Fig. 1: Einführung einer DSB-Erkennungssequenz in das Plastom mittels Doppeltem "Cross-Over"

In einer besonders bevorzugten Ausführungsform 1 wird zunächst ein DSBE-Konstrukt in Plastiden einer höheren Pflanze eingebracht. Das DSBE-Konstrukt wird in dieser Ausführungsform bevorzugt mit homologen Zielregionen und mit einem exprimierbaren Selektionsmarker (Promotor - 5'UTR - Selektionsmarker - 3'UTR) ausgerüstet und enthält in dieser Ausführungsform bevorzugt eine

## 81

Erkennungsregion für ein DSBI-Enzym, welches bevorzugt keine natürliche Erkennungssequenz im Plastidengenom der (nicht-transformierten) betrachteten Pflanze besitzt. Das DSBE-Konstrukt kann - optional - bereits weitere Gene von Interesse kodieren. Es werden überwiegend homoplastome Masterpflanzen erzeugt (Fig. 1).

2. Fig. 2A-E: Einführung einer Insertionssequenz mit einer Expressionskassette für eine DSBI-Enzym und ggf. Selektionsmarker sowie weiteren Genen von Interesse

10

Explantate der durch Ausführungsform 1 hergestellten Masterpflanzen werden für eine weitere Transformation mit einem erfundungsgemäßen Transformationskonstrukt genutzt. Das erfundungsgemäße Transformationskonstrukt besitzt bevorzugt beidseitig

15 (Fig. 2A, 2B) oder einseitig (Fig. 2C, 2D) der Insertionssequenz Bereiche, die homolog zum den Sequenzen umgebend die Insertionsstelle des DSBE-Konstruktes sind. Die Insertion erfolgt dann über homologe Rekombination (z.B. Cross-Over) bzw. über eine Reperatursynthese.

20

Besonders bevorzugt sind die zu insertierenden Sequenzen - nach innen an die Homologiesequenzen anschließend - von Teilen der DSB-Erkennungssequenz flankiert (nDS), die den in Folge eines Schnittes mit dem DSBI-Enzym entstehenden Teilen, entsprechen

25 (Fig. 2B). Somit enthält die Insertionssequenz Sequenzen, die den in Folge eines Schnittes im Plastom entstehenden Enden unmittelbar entsprechenden und so einen besonders effizienten Einbau gewährleisten.

30 Besitzt das Transformationskonstrukt bzw. die Insertionssequenz keine solche homologen Bereiche, wird die Insertionssequenz an diesen Enden bevorzugt mit überhängenden Enden versehen, die auch von dem DSBI-Enzym nach Schneiden des Plastoms der Masterpflanze erzeugt werden (Fig. 2E).

35

Liegt nur eine Homologiesequenz vor, so grenzt an diese in einer besonders bevorzugten Ausführungsform eine nDS-Sequenz (s.o. wie für Fig. 2B beschrieben) an, während die andere Seite der Insertionssequenz mit überhängenden Enden versehen ist, die den

40 vom DSBI-Enzym im Plastom der Masterpflanze erzeugten entsprechen (Fig. 2D).

Optional kodiert die Insertionssequenz für einen weiteren, exprimierbaren Selektionsmarker, der sich funktionell von dem 45 des DSBE-Konstruktes unterscheidet, ggf. ein oder mehrere exprimierbare Gene von Interesse und das exprimierbare DSBI-Enzym, welches die durch das DSBE-Konstrukt eingebrachte

Erkennungssequenz an der Insertionsstelle im Plastidengenom der Masterpflanze schneidet. Die Insertion der Insertionssequenz des Transformationskonstruktes erfolgt dabei an einer Stelle derart, dass besagte Erkennungsregion nach der Insertion nicht mehr 5 funktionell ist.

3. Fig.3: Einführung einer DSB-Erkennungssequenz in das Plastom mittels Doppeltem "Cross-Over" in einen transkriptionell aktiven Bereich

10

In einer weiteren, besonders bevorzugten Ausführungsform 2 wird zunächst ein DSBE-Konstrukt in Plastiden einer höheren Pflanze eingebracht. Das DSBE-Konstrukt wird in dieser Ausführungsform bevorzugt mit homologen Zielregionen und mit einem exprimierbaren 15 Selektionsmarker ausgerüstet, wobei ein endogener Promotor des Plastoms dazu genutzt wird, und enthält zusätzlich bevorzugt eine Erkennungsregion für ein DSBI-Enzym, welches bevorzugt keine natürliche Erkennungssequenz im Plastidengenom der (nicht-transformierten) betrachteten Pflanze besitzt. Das DSBE-Konstrukt kann 20 bereits für Gene von Interesse kodieren. Es werden überwiegend homoplastome Masterpflanzen erzeugt (Fig. 3).

4. Fig.4A-E: Einführung einer Insertionssequenz mit einer Kassette kodierend für eine DSBI-Enzym und ggf. Selektions- 25 marker sowie weiteren Genen von Interesse

Explantate der durch Ausführungsform 2 hergestellten Masterpflanzen werden für eine weitere Transformation mit einem erfindungsgemäßen Transformationskonstrukt genutzt. Das erfindungsgemäße Transformationskonstrukt besitzt bevorzugt beidseitig 30 (Fig. 4A, 4B) oder einseitig (Fig. 4C, 4D) der Insertionssequenz Bereiche, die homolog zum den Sequenzen umgebend die Insertionsstelle des DSBE-Konstruktes sind. Die Insertion erfolgt dann über homologe Rekombination (z.B. Cross-Over) bzw. eine Reperatur- 35 synthese. Besonders bevorzugt sind die zu insertierenden Sequenzen - nach innen an die Homologiesequenzen anschließend - von Teilen der DSB-Erkennungssequenz flankiert (nDS), die den in Folge eines Schnittes mit dem DSBI-Enzym entstehenden Teilen, entsprechen (Fig. 4B). Somit enthält die Insertionssequenz 40 Sequenzen, die den in Folge eines Schnittes im Plastom entstehenden Enden unmittelbar entsprechenden und so einen besonders effizienten Einbau gewährleisten.

Besitzt das Transformationskonstrukt bzw. die Insertionssequenz 45 keine solche homologen Bereiche, wird die Insertionssequenz an diesen Enden bevorzugt mit überhängenden Enden versehen, die auch

von dem DSBI-Enzym nach Schneiden des Plastoms der Masterpflanze erzeugt werden (Fig. 4E).

Liegt nur eine Homologiesequenz vor, so grenzt an diese in einer 5 besonders bevorzugten Ausführungsform eine nDS-Sequenz (s.o. wie für Fig. 4B beschrieben) an, während die andere Seite der Insertionssequenz mit überhängenden Enden versehen ist, die den vom DSBI-Enzym im Plastom der Masterpflanze erzeugten entsprechen (Fig. 4D).

10

Optional kodiert die Insertionssequenz für einen weiteren, exprimierbaren Selektionsmarker, der sich funktionell von dem des DSBE-Konstruktes unterscheidet, ggf. ein oder mehrere exprimierbare Gene von Interesse und das exprimierbare DSBI- 15 Enzym, welches die durch das DSBE-Konstrukt eingebrachte Erkennungssequenz an der Insertionsstelle im Plastidengenom der Masterpflanze schneidet. Die Insertion der Insertionssequenz des Transformationskonstruktes erfolgt dabei an einer Stelle derart, dass besagte Erkennungsregion nach der Insertion nicht 20 mehr funktionell ist.

5. Fig. 5A-E: Einführung einer Insertionssequenz mit einer Kassette 25 kodierend für eine DSBI-Enzym und ggf. Selektionsmarker sowie weiteren Genen von Interesse unter Nutzung natürlicher, endogener DSB-Erkennungssequenzen

In einer weiteren, ganz besonders bevorzugten Ausführungsform 3 enthält ein erfindungsgemäßes Transformationskonstrukt ein 30 exprimierbares DSBI-Enzym, welches eine endogene, natürliche Erkennungssequenz im Plastom der betrachteten Pflanze besitzt.

Explantate dieser natürlichen Masterpflanzen werden für eine Transformation mit einem erfindungsgemäßen Transformations- 35 konstrukt genutzt. Das erfindungsgemäße Transformationskonstrukt besitzt bevorzugt beidseitig (Fig. 5A, 5B) oder einseitig (Fig. 5C, 5D) der Insertionssequenz Bereiche, die homolog zum den Sequenzen umgebend die Insertionsstelle des DSBE-Konstruktes sind. Die Insertion erfolgt dann über homologe Rekombination 40 (z.B. Cross-Over) bzw. Reperatursynthese.

Besonders bevorzugt sind die zu insertierenden Sequenzen - nach innen an die Homologiesequenzen anschließend - von Teilen der DSB-Erkennungssequenz flankiert (nDS), die den in Folge eines 45 Schnittes mit dem DSBI-Enzym entstehenden Teilen, entsprechen (Fig. 5B). Somit enthält die Insertionssequenz Sequenzen, die den in Folge eines Schnittes im Plastom entstehenden Enden unmittel-

bar entsprechenden und so einen besonders effizienten Einbau gewährleisten.

Besitzt das Transformationskonstrukt bzw. die Insertionssequenz 5 keine solche homologen Bereiche, wird die Insertionssequenz an diesen Enden bevorzugt mit überhängenden Enden versehen, die auch von dem DSBI-Enzym nach Schneiden des Plastoms der Masterpflanze erzeugt werden (Fig. 5E).

10 Liegt nur eine Homologiesequenz vor, so grenzt an diese in einer besonders bevorzugten Ausführungsform eine nDS-Sequenz (s.o. wie für Fig. 5B beschrieben) an, während die andere Seite der Insertionssequenz mit überhängenden Enden versehen ist, die den vom DSBI-Enzym im Plastom der Masterpflanze erzeugten entsprechen 15 (Fig. 5D).

Die Insertion der Insertionssequenz des Transformationskonstruktes erfolgt dabei bevorzugt an einer Stelle derart, dass besagte Erkennungsregion nach der Insertion nicht mehr 20 funktionell ist. Die Insertionssequenz kodiert bevorzugt für einen exprimierbaren Selektionsmarker ( $S'$ ), ein oder mehreren Genen von Interesse sowie für das exprimierbare DSBI-Enzym. Der Selektionsmarker ist optional.

25 6. Fig. 6A-E: Einführung einer Insertionssequenz mit einer Kassette kodierend für Gene von Interesse und ggf. Selektionsmarker und Einbringung eines DSBI-Enzyms in trans

In weiteren bevorzugten Ausführungsformen 4 wird das DSBI Enzym 30 nicht von dem Transformationskonstrukt kodiert, sondern entweder in trans exprimiert (in Plastiden oder als PLS-Fusionsprotein im Kern) bzw. in Form von RNA oder als Protein in die Plastiden transfiziert. Das DSBI-Enzym erkennt entweder eine künstlich eingeführte (Fig. 6A, 6B) oder natürliche (Fig. 6C, 6D) DSB- 35 Erkennungssequenz. Diese Ausführungsform ist insbesondere dann bevorzugt, wenn das Transformationskonstrukt keine Promotor-elemente umfasst, und eine Expression der kodierten Gene erst nach Insertion in das Plastom unter Verwendung plastidärer, endogener Promotoren realisiert wird.

40 Wie bei den bereits oben beschriebenen Ausführungsbeispielen besitzt das Transformationskonstrukt bevorzugt beidseitig (Fig. 6A, 6B) oder einseitig (nicht gezeigt) der Insertionssequenz Bereiche, die homolog zum den Sequenzen umgebend die 45 Insertionsstelle des DSBE-Konstrukt sind. Die Insertion erfolgt

85

dann über homologe Rekombination (z.B. Cross-Over) bzw.  
Reperatursynthese.

Besonders bevorzugt sind die zu insertierenden Sequenzen - nach  
5 innen an die Homologiesequenzen anschließend - von Teilen der  
DSB-Erkennungssequenz flankiert (nDS), die den in Folge eines  
Schnittes mit dem DSBI-Enzym entstehenden Teilen entsprechen  
(Fig. 6B, 6D). Somit enthält die Insertionssequenz Sequenzen,  
die den in Folge eines Schnittes im Plastom entstehenden Enden  
10 unmittelbar entsprechenden und so einen besonders effizienten  
Einbau gewährleisten.

Besitzt das Transformationskonstrukt bzw. die Insertionssequenz  
keine solche homologen Bereiche, wird die Insertionssequenz an  
15 diesen Enden bevorzugt mit überhängenden Enden versehen, die auch  
von dem DSBI-Enzym nach Schneiden des Plastoms der Masterpflanze  
erzeugt wird (Fig. 6E). Das Transformationskonstrukt kann zusätz-  
lich eine Sequenz kodierend für ein DSBI-Enzym umfassen. Die  
Expression erfolgt jedoch erst nach erfolgreicher Insertion in  
20 das Plastom, so dass die Bereitstellung einer ersten Menge an  
funktioneller RNA oder Protein eines DSBI-Enzyms wünschenswert  
ist.

Liegt nur eine Homologiesequenz vor, so grenzt an diese in einer  
25 besonders bevorzugten Ausführungsform eine nDS-Sequenz (s.o. wie  
für Fig. 6B, 6D beschrieben) an, während die andere Seite der  
Insertionssequenz mit überhängenden Enden versehen ist, die den  
vom DSBI-Enzym im Plastom der Masterpflanze erzeugten entsprechen  
(nicht gezeigt).

30.

Die Insertion der Insertionssequenz des Transformations-  
konstruktes erfolgt dabei bevorzugt an einer Stelle derart,  
dass besagte Erkennungsregion nach der Insertion nicht mehr  
funktionell ist.

35

7. Fig. 7A-E: Einführung einer Insertionssequenz umfassend einer  
Intronsequenz mit einer Kassette kodierend für Gene  
von Interesse und ggf. Selektionsmarker bzw. DSBI-  
Enzyme

40

In einer weiteren, ganz besonders bevorzugten Ausführungsform 5  
ist das Gen von Interesse (sowie optional ein Selektionsmarker S'  
und/oder das DSBI-Enzym) innerhalb eines Introns kodiert, welches  
an der gewählten Insertionsstelle funktionell ist, d.h. aus dem  
45 dort gebildeten Transkript spleißen kann.

## 86.

Das erfindungsgemäße Transformationskonstrukt besitzt bevorzugt beidseitig (Fig. 7A; 7B) oder einseitig (nicht gezeigt) der Insertionssequenz Bereiche, die homolog zu den Sequenzen umgebend die Insertionsstelle des DSBE-Konstruktes sind. Die Insertion 5 erfolgt dann über homologe Rekombination (z.B. Cross-Over) bzw. Reparatursynthese.

Besitzt das Transformationskonstrukt bzw. die Insertionssequenz keine solche homologen Bereiche, wird die Insertionssequenz an 10 diesen Enden bevorzugt mit überhängenden Enden versehen, die auch von dem DSBI-Enzym nach Schneiden des Plastoms der Masterpflanze erzeugt wird (nicht gezeigt).

Die Kontrolle der Expression kann mittels eines auf dem Transformationskonstrukt enthaltenen Promotors (Fig. 7B) oder einem 15 endogenen, plastidären Promotor realisiert werden (Fig. 7A). Im ersten Fall ist das DSBI-Enzym bevorzugt auf dem Transformationskonstrukt enthalten (Fig. 7B), während es im letzteren Fall (zumindest parallel) entweder in trans exprimiert (in Plastiden 20 oder als PLS-Fusionsprotein im Kern) bzw. in Form von RNA oder als Protein in die Plastiden transfiziert wird (Fig. 7A).

Die Insertion der Insertionssequenz des Transformationskonstruktes erfolgt dabei bevorzugt an einer Stelle derart, 25 dass besagte Erkennungsregion nach der Insertion nicht mehr funktionell ist. Das Transformationskonstrukt kann zusätzlich optional eine Sequenz kodierend für ein DSBI-Enzym umfassen. Die Expression erfolgt jedoch erst nach erfolgreicher Insertion in das Plastom, so dass die Bereitstellung einer ersten Mengen an 30 funktioneller RNA oder Protein eines DSBI-Enzyms wünschenswert ist.

8. Fig. 8: Deletion von Seuenzen mittels intramolekularer homologer Rekombination induziert durch sequenz- 35 spezifische Doppelstrangbrüche

Sequenzen - beispielsweise kodierend für Selektionsmarker bzw. DSBI-Enzyme - sind bei allen oben beschriebenen Ausführungsformen bevorzugt von Homologiesequenzen H1 und H2 flankiert, die 40 eine ausreichende Länge und Homologie aufweisen, um miteinander zu rekombinieren. Die Rekombination wird durch die Induktion mindestens einen Doppelstrangbruches zwischen den beiden Homologiesequenzen gelegenen DSB-Erkennungssequenz induziert. Zur Induktion des Doppelstrangbruches wird bevorzugt ein DSBI-Enzym 45 transient exprimiert oder eingebracht (Fig. 8).

Dem Fachmann ist bewusst, dass die Reihenfolge der exprimierten Gene in einem Operon austauschbar ist und insofern in den oben beschriebenen Ausführungsformen variieren kann. Auch kann bei Verwendung nur einer Homologiesequenz zu Insertion 5 der Insertionssequenz diese an der 5'- (wie in den Abbildungen gezeigt) oder 3'-Seite des Doppelstrangbruches lokalisiert sein. Prinzipiell kann das DSBI-Enzym auf dem Transformationskonstrukt und/oder separat (im Kern oder Plastiden) exprimiert oder andersweitig - beispielsweise durch Transfektion mit RNA oder Protein 10 in Plastiden - eingebracht werden.

9. Fig.9: Southern Analyse von überwiegend homotransplastomen Pflanzen

15 Wildtyp und überwiegend homotransplastome Masterpflanzen wurden bezüglich der Modifikation (Einführung einer DSB-Erkennungssequenz) analysiert (vgl. Beispiel 4). Durch die Modifikation wurde eine Bande von 1750 bp detektiert (Bahn 2, 3, und 4 entsprechend den Linien CB199NTH-4, -6, und -8), während in der nicht 20 modifizierten Wildtyppflanze eine Bande von 3100 bp detektiert wurde (Bahn 1).

10. Fig.10: Modifizieren der IGS des Tetrahymena LSU Intron.

25 Großbuchstaben zeigen die Sequenz des Intron, während Kleinbuchstaben die Sequenz der umgebenden Exons darstellen. Gezeigt sind die flankierenden Exonsequenzen, der 5'- und 3'- Anteil des Introns bzw. des Intronderivates sowie die Sequenz umfassend die IGS. Balken zwischen den Basen deuten mögliche Basenpaarungen an, 30 die zur Initiation des Spleißvorganges ausgebildet werden können.

A: Gezeigt sind die genannten Sequenzabschnitte des natürlicherweise vorkommenden Tetrahymena LSU Intron in seiner natürlichen Exon-Umgebung.

35

B: Gezeigt sind die genannten Sequenzabschnitte des im Rahmen dieser Erfindung erstellten Tetrahymena LSU Intron-Derivat (TetIVS2a) in der vordefinierten Exonumgebung, wie sie beim CpLSU5 Intron innerhalb der DSB-Erkennungssequenz des DSBI-Enzyms I-CpaI zu finden ist. Fettgedruckte Buchstaben repräsentieren die im Vergleich zur natürlichen Sequenz vorgenommenen Mutationen.

## 11. Fig.11:

- A: Southern Analyse mit von mit BamHI geschnittener Gesamt-DNA aus den Tabaklinien CB255+435NTTH-16b, -16c, -19 und -20. Als Sonde wurde ein Bereich der 16SrDNA genutzt. Die Banden, welche Plastomkopien repräsentieren, die dem Wildtyp (WT) entsprechen (ca. 3,2 kb Bande detektiert) und die das Transgen tragen (TG) (ca. 2,3 kb Bande detektiert) sind durch Pfeile gekennzeichnet.
- B: Schematische Dartstellung von transplastomen Tabakpflanzen, die durch die Insertion der Insertionssequenz aus pCB435-45 entstanden sind; sowie die bei einer entsprechenden Southern-Analyse (vgl. A) zu erwartenden Banden.
- C: Southern Analyse mit von mit HindIII geschnittener Gesamt-DNA aus den Tabaklinien CB255+435NTTH-16b, -16c, -19 und -20. Als Sonde wurde ein Bereich der 23SrDNA genutzt. Die Banden, welche Plastomkopien repräsentieren, die dem Wildtyp (WT) entsprechen (ca. 1,1 kb Bande detektiert) und die das Transgen (TG) tragen (ca. 1,5 kb Bande detektiert) sind durch Pfeile gekennzeichnet.
- D: Schematische Dartstellung von transplastomen Tabakpflanzen, die durch die Insertion der Insertionssequenz aus pCB255-1 entstanden sind; sowie die bei einer entsprechenden Southern-Analyse (vgl. C) zu erwartenden Banden.

## 12. Fig. 12:

- A: Wildtyp und überwiegend homotransplastome Masterpflanzen wurden bezüglich der Modifikation (Einführung einer der I-PpoI DSB-Erkennungssequenz) in einem Southern analysiert. (vgl. Beispiel 14.2). Durch die Modifikation wurde in der hier mit EcoRI behandelten DNA eine Bande von etwa 1,7 kb detektiert (TG) (Bahn 1 und 4 entsprechend den Linien CB456NTH-1 und -15), während in der nicht modifizierten Wildtyppflanze (WT) eine Bande von etwa 3,1 kb detektiert wurde (Bahn 6). (wt - nicht modifizierte Wildtyppflanze; wildtype - zeigt die erwartete Fragmentgröße in nicht modifizierte Wildtyppflanzen an; transgenic - zeigt die erwartete Fragmentgröße in Pflanzen CB456NTH an).
- B: Schematische Darstellung des EcoRI Fragments, welches in A durch die Hybridisierung mit der Sonde (probe) in einer modifizierten Pflanze CB456NTH zu erwarten war. (trnV - Gen codierend für eine tRNA-Val; rrn16 - Gen codierend für die

16SrRNA; aadA - Gen codierend für einen Selektionsmarker;  
3'psbA (Synec) - nicht kodierender Bereich stromabwärts des  
psbA-1 Gens aus Synechocystis, hier eingebaut in die Express-  
5 sionskassette für den Selektionsmarker aadA; Psynth. - syn-  
theticischer Promotor abgeleitet von der Konsensussequenz für  
E. coli σ70 Promotoren; DSB-E: DSB-Erkennungssequenz).

### Beispiele

#### 10 Allgemeine Methoden:

Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die 15 im Rahmen der vorliegenden Erfindung durchgeführten Klonierungs- schritte wie z.B. Restriktionsspaltungen, Agarosegelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nuklein- säuren auf Nitrozellulose und Nylonmembranen, Verknüpfen von DNA-Fragmenten, Transformation von E. coli Zellen, Anzucht von 20 Bakterien, Vermehrung von Phagen und Sequenzanalyse rekombinanter DNA werden wie bei Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6 beschrieben durchgeführt. Die Sequenzierung rekombinanter DNA-Moleküle erfolgt mit einem 25 Laserfluoreszenz-DNA-Sequenzierer ALF-Express (Pharmacia, Upsala, Schweden) nach der Methode von Sanger (Sanger et al. (1977) Proc Natl Acad Sci USA 74:5463-5467)..

#### Beispiel 1: Erstellen eines Basis Vektors für die Plastiden- transformation

30 Zunächst wurden aus dem Plastom der Tabak Varietät SR1 die ausgewählten Zielregionen mittels PCR kloniert. Die linke Zielregion wurde dabei mit den Primern p19 und p20 amplifiziert.

35 p19: 5'-TAAGGCCCTCGTAGAACGG-3' (SEQ ID NO: 16)

p20: 5'-GGGGTACCAAATCCAATG-3' (SEQ ID NO: 17)

Zur Amplifikation der rechten Zielregion wurden die Primer p21 40 und p22 verwendet, wobei durch letzteren Primer in den amplifizierten Teil der 16SrDNA zusätzlich zur SR1 Resistenz (binding type marker) auch noch eine Spectinomycin-Resistenz eingebracht wurde.

45 p21: 5'-GGAGCTCGCTCCCCCGCCGTGTTTC-3' (SEQ ID NO: 18)

90

p22: 5'-GATGCATGATGACTTGACGGCATCCTC-3' (SEQ ID NO: 19)

Die beiden amplifizierten Regionen wurden in pBluescript bzw. pZeroBlunt kloniert und sequenziert. Die linke und rechte Zielregionen wurden anschließend in das Rückgrat des pUC19 Vektors kloniert. Dazu wurden die Schnittstellen Eco0109I und PvUII des Vektors genutzt. Zwischen die linke und rechte Zielregion wurde eine multiple Klonierungsstelle aus dem pBluescript (von KpnI bis SacI) kloniert. Diese befindet sich im Basisvektor für die Plastidentransformation zwischen den beiden plastidär kodierten Genen trnV und rrn16. Dieser Basisvektor für die Plastidentransformation erhielt die Bezeichnung pCB42-94 (SEQ ID NO: 1). Der Vektor enthält nachfolgende Sequenzelemente:

- 15 a) Position komplementär bp 55-1405: Rechte Zielregion mit dem partiellen Gen der 16SrRNA (komplementär bp 56 bis 1322). In letzterem sind Mutationen für die Streptomycin-Resistenz (SR1, Position bp 346) und Spectinomycin-Resistenz (SPC1, Position bp 68).
- 20 b) Position komplementär bp 2374 bis 1510: Linke Zielregion umfassend u.a. ORF131 (bp 1729 bis 2124) und trnV Gen (komplementär bp 1613 bis 1542).
- c) Position bp 1404 bis 1511: multiple Klonierungsstelle
- 25 d) Position bp 2629 bis 3417: Ampicillin Resistenz im Vektorrückgrat

Beispiel 2: Erstellen eines Vektors (pCB199-3) für die Einbringung einer nicht natürlicherweise vorkommenden Erkennungsregion für die Homing Endonukleasen I-PpoI in das Plastom von Tabak

In die multiple Klonierungsstelle des Basisvektors pCB42-94 (SEQ ID NO: 1) für die Plastidentransformation wurden sukzessive verschiedene Elemente kloniert:

- a) frt Erkennungsregion (mutiert, enthält keine XbaI Schnittstelle; komplementär 1307-1354)
- 40 b) Expressionskassette zur Expression des Markergens aadA bestehend aus:
- i) Promotor des Gens für 16SrRNA (komplementär 1191-1281)

45

91

- ii) 5' untranslatierte Bereich des Tabak rbcL Gens (komplementär 1167-1184) einschließlich mutierter 5' Anteilen des rbcL Gens (Duplikation von 6 AS des rbcL Gens, teilweise mutiert, als Folge der Klonierungsstrategie, so dass eine Fusion codierend für insgesamt 12 Aminosäuren (komplementär 1131-1166) mit dem nachfolgendem Element, dem aadA Gen, entstand)
- 10 iii) aadA Gen (komplementär 336-1130)
- iv) der 3' Bereich des psbA Gens (komplementär 232-323)
- 15 c) Core-Erkennungsregion für die Homing Endonuklease I-PpoI (komplementär 176-190).

15 Anstelle der multiplen Klonierungsstelle im Basisvektor für die Plastidentransformation enthält dieser Vektor mit der Bezeichnung pCB199-3 die aufgeführten Elemente im Rahmen der Nukleinsäuresequenz mit der SEQ ID NO: 2. Angegeben ist die Sequenz, die 20 die komplette MCS von KpnI bis SacI ersetzt. In der angegebenen Sequenz gibt es aufgrund der Klonierungsstrategie jedoch keine KpnI Schnittstelle mehr.

25 Beispiel 3: Erstellen eines weiteren Vektors (pCB401-20) für die Einbringung einer nicht natürlicherweise vorkommende Erkennungsregion für die Homing Endonukleasen I-PpoI in das Plastom von Tabak

30 Im Gegensatz zu dem in Beispiel 2 beschriebenen Vektor pCB199-3 enthält der hier beschriebene Vektor keinen Promotor und 3'UTR, der direkt mit dem Selektionsmarker aadA verknüpft wurde. Vielmehr wird die Expression des aadA Gens ausgehend von dem Promotor des trnV Gens, welches im Plastidengenom bzw. in der linken Zielregion stromaufwärts des aadA Gens lokalisiert ist, 35 gesteuert. Ziel bei der Erstellung dieses Vektors war es, Sequenzduplikationen durch das Ausnutzen von regulatorischen Bereichen aus dem Tabakplastidengenom zu vermeiden. Dazu wurden in die multiple Klonierungsstelle des Basisvektors pCB42-94 für die Plastidentransformation sukzessive verschiedene Elemente 40 kloniert:

- a) Ribosomenbindestelle (komplementär bp 1033 bis 1050)
- b) aadA Gen (komplementär bp 238 bis 1032)

45

- c) Core-Erkennungsregion für die Homing Endonukleasen I-PpoI  
(komplementär bp 176 bis 190)

Der so erhaltene Vektor vermittelt auch in *E. coli* eine Spectino-  
5 mycin Resistenz. Anstelle der multiplen Klonierungsstelle im  
Basisvektor pCB42-94 (SEQ ID NO: 1) für die Plastidentrans-  
formation enthält dieser Vektor mit der Bezeichnung pCB401-20  
die aufgeführten Elemente im Rahmen der Nukleinsäuresequenz mit  
der SEQ ID NO: 3. Auch hier ist die gesamte die MCS ersetzende  
10 Sequenz (von SacI bis KpnI) angegeben.

Beispiel 4: Erstellen von überwiegend homoplastomischen Tabak-  
Masterpflanzen, die eine nicht natürliche DSB-  
Erkennungssequenz enthalten

15

Das Plasmid pCB199-3 wurde in die Plastiden von Tabak (*Nicotiana tabacum* cv. Petit Havana) wie im folgenden beschrieben eingebracht. Die regenerierten Pflanzen wurden CB199NTH genannt. Unabhängige Linien wurden mit verschiedenen Endnummern (z.B.  
20 CB199NTH-4) versehen.

Analog wird der Vektor pCB401-20 in die Plastiden von Tabak eingebracht. Die resultierenden Pflanzen werden entsprechend mit CB401NTH bezeichnet.

25

Zunächst wurden von *in vitro* gewachsenen Pflanzen mit einem sterilen Korkbohrer Blattscheibchen mit einem Durchmesser von 2,0 bis 2,5 cm ausgestochen und je mit der Blattoberseite auf eine Petrischale mit Beschussmedium [MS-Salze (Sigma-Aldrich):  
30 4,3 g/L; Saccharose: 30,0 g/L, Phytoagar (Duchefa, P1003): 0,6 % (w/v); pH 5,8; nach dem Autoklavieren wurden 1,0 mg/L Thiamin (Duchefa, T0614) und 0,1 g/L myo-Inositol (Duchefa, I0609) zugegeben] platziert. Die vom Agar abgewandte Blattunterseite wurde mittels der Partikelkanone anschließend beschossen. Dazu wurde  
35 zunächst die zu transformierende Plasmid-DNA (aus *E. coli* gereinigt mittels Nucleobond AX100 Macherey & Nagel) nach folgendem Protokoll auf 0,6 µm große Goldpartikel aufgebracht ("Coating"). Zunächst wurden 30 mg Goldpulver (BioRad) in Ethanol aufgenommen. 60 µl von der Goldsuspension wurden in ein neues  
40 Eppendorfgefäß überführt und die Goldpartikel durch Zentrifugation (für 10 Sekunden) sedimentiert. Die Goldpartikel wurden zweimal in je 200 µl sterilem Wasser gewaschen und nach einem weiteren Zentrifugationsschritt in 55 µl Wasser aufgenommen. Unter ständigem Mischen (Vortexen) wurden schnell hinzugegeben:  
Unter ständigem Mischen (Vortexen) wurden schnell hinzugegeben:

45

- 5 µl Plasmid-DNA (1 µg/µl)
- 50 µl 2,5 M CaCl<sub>2</sub>

- 20 µl 0,1 M Spermidine

Die Suspension wurde sodann für weitere 3 Minuten auf dem Vortex gemischt und anschließend kurz zentrifugiert. Die sedimentierten Gold-DNA-Komplexe wurden ein- bis zweimal in je 200 µl Ethanol gewaschen und nach einem weiteren Zentrifugationsschritt schließlich in 63 µl Ethanol aufgenommen. Pro Schuss wurden 3,5 µl (entsprechend 100 µg Gold) dieser Suspension auf einen Macrocarrier aufgebracht.

10

Gemäß den Herstellerangaben wurde die Partikelkanone (BioRad, PDS1000He) vorbereitet und die Blattexplantate in einem Abstand von 10 cm mit den Gold-DNA-Komplexen beschossen. Dabei wurden folgende Parameter verwendet: Vakuum: 27 inch Hg, Druck 15 1100 psi. Die Explantate werden nach dem Beschuss für 2 Tage in Klimakammern (24°C, 16 h Licht, 8 h Dunkelheit) inkubiert und anschließend mit einem Skalpell in etwa 0,5 cm<sup>2</sup> große Segmente zerteilt. Diese Segmente wurden dann auf Regenerationsmedium [Beschussmedium zuzüglich 1 mg/L 6-Benzylaminopurin (BAP, Duchefa, 20 B0904) und 0,1 mg/L Naphthylessigsäure (NAA, Duchefa, N0903)] zuzüglich 500 mg/L Spectinomycin (Duchefa, S0188) überführt und für 10 bis 14 Tage unter oben genannten Bedingungen in einer Klimakammer inkubiert. Nach Ablauf dieser Zeit wurden die Blattsegmente auf frisches Regenerationsmedium zuzüglich 500 mg/L Spectinomycin überführt. Dieser Vorgang wurde so lange wiederholt, bis sich grüne Sprosse an den Explantaten bildeten. Die Sprosse wurden mit einem Skalpel abgetrennt und auf Anzuchtmedium (wie Beschussmedium, jedoch 10 g/L Saccharose anstelle von 30 g/L Saccharose) zuzüglich 500 mg/L Spectinomycin angezogen.

30

Optional können, um möglichst überwiegend homoplastome Pflanzen zu erhalten, von den regenerierten Pflanzen wiederum Explantate abgeschnitten und auf Regenerationsmedium mit 1000 mg/L Spectinomycin gelegt werden. Regenerierende Sprosse werden in Gläser mit 35 Anzuchtmedium zuzüglich 500 bis 1000 mg/L Spectinomycin gesetzt. Nach dem Bewurzeln werden die Pflanzen ins Gewächshaus transferiert und dort auf Erde bis zur Samenreife herangezogen.

Im Fall der Transformation mit dem Plasmid pCB199-3 wurden 40 8 Pflanzen erhalten, die Resistenz gegenüber Spectinomycin zeigten. Mittels PCR und Southern-Analysen wurde nachgewiesen, dass drei dieser Linien (Linien CB199NTH-4, -6 und -8) auch tatsächlich das aadA Gen in das Plastidengenom eingebaut haben.

45 Zur Analyse der transplastomen Pflanzen nach Southern wurde Gesamt-DNA aus Blättern von transformierten und nicht transformierten Pflanzen mit Hilfe des GenElute Plant Genomic DNA Kit

(Sigma) isoliert. Die DNA wurde dabei in 200 µl Eluat aufgenommen. 86 µl davon wurden je mit 10 µl 10x Restriktionspuffer sowie 40 U Restriktionsendonuklease versetzt und 4 bis 8 h bei der für das Restriktionsenzym empfohlenen Temperatur inkubiert.

5 Anschließend wurde die DNA in dem Fachmann bekannter Weise mit Ethanol präzipitiert und anschließend in 20 µl Wasser aufgenommen. Die Proben wurden anschließend gemäß dem Fachmann bekannten Verfahren auf einem Agarosegel aufgetrennt, die DNA in dem Gel denaturiert und mittels Kapillarblot auf Nylonmembran

10 übertragen.

Eine entsprechende Sonde für die radioaktive Hybridisierung wurde mit Hilfe des HighPrime (Roche) Systems erstellt. Die Membran wurde zunächst 1 h bei 65°C mit HybPuffer (1 % (w/v)) Rinderseumalbumin; 7 % (w/v) SDS; 1 mM EDTA; 0,5 M Natriumphosphat-Puffer, pH 7,2) vorhybridisiert. Anschließend wurde die hitzenaturierte Sonde hinzugegeben, und es wurde über Nacht bei 65°C hybridisiert. Die Blots wurden anschließend wie folgt gewaschen: einmal Spülen mit 2 x SSPE/0,1 % SDS; für 15 min.

20 bei 65°C waschen mit 2 x SSPE/0,1 % SDS; für 15 min. bei 65°C waschen mit 1 x SSPE/0,1 % SDS und ggf. wurde der letzte Schritt noch einmal wiederholt (20 x SSPE ist 3 M NaCl; 0,2 M NaH<sub>2</sub>PO<sub>4</sub>; 0,5 M EDTA; pH 7,4).

25 Die Hybridisierung wurde anschließend mit Hilfe eines Phospho-imagers (Molecular Imager FX, BioRad) analysiert.

Beispielsweise wurde mit PstI geschnittene Gesamt-DNA verschiedener Pflanzen, die nach der Transformation mit pCB199-3 regeneriert wurden, mit dem aadA Gen als radioaktiv markierte Sonde (793bp PstI\NcoI Fragment aus pCB199-3) hybridisiert. Dabei konnte festgestellt werden, dass die Linien CB199NTH-4, -6 und -8 tatsächlich das aadA Gen in die DNA eingebaut hatten. Des weiteren wurde mit EcoRI und XbaI geschnittene gesamt DNA von CB199NTH-4, -6 und -8 mit einer radioaktiv markierten Sonde (1082 bp Bsp120I/SacI Fragment aus pCB199-3) hybridisiert, welche mit einem Teil der 16S rDNA hybridisiert. Während im Wildtyp (nicht transformierte Pflanze) erwartungsgemäß eine Bande von etwa 3100 bp nachgewiesen wurde, wurde in den transplastomatischen Linien überwiegend eine Bande bei 1750 bp detektiert, die sich aufgrund der Insertion der Insertionssequenz aus pCB199-3 in das Plastom ergibt (Fig. 9). Die erhaltenen Pflanzen können als überwiegend homotransplastom verstanden werden.

Beispiel 5: Erstellen von Transformationsvektoren, die zur Transformation der Plastiden von Masterpflanzen CB199NTH mittels des artifiziellen Homing-Prozesses genutzt werden können

5

### 5.1 Klonieren der Homing Endonuklease I-PpoI

Die Homing Endonuklease I-PpoI wurde aus 26 synthetisch erstellten Oligonukleotiden mittels PCR in Anlehnung an die 10 Methode von Stemmer WPC et al. (1995) Gene 164: 49-53 erstellt (SEQ ID NO: 11). Die zugrunde liegende Sequenz wurde von der veröffentlichten Sequenz abgeleitet (Accession No. M38131 Nukleotide 86 bis 577). Dabei wurden einige Mutationen eingeführt, um Erkennungsstellen für Restriktionsendonukleasen aus 15 dem Gen zu entfernen, von denen aber keine eine veränderte Aminosäuresequenz bedingt. Anschließend wurden sukzessiv folgende Elemente in einem pBluescript KS (Stratagene) Vektorrückgrat vereint, um eine I-PpoI Expressionskassette zu erstellen. Die Sequenz ist von den Schnittstellen KpnI und SacI flankiert.

20

- a) Position 21 bis 111: Prrn Promotor
- b) Position 118 bis 135: 5'-untranslatierte Region des rbcL Gens gefolgt von 18 bp kodierend für 6 Aminosäuren des 25 rbcL-Protein (bp 136-152)
- c) Position 154 bis 645: Nukleinsäuresequenz kodierend für I-PpoI.

30 d) Position 688-779: 3'-untranslatierte Region des psbA-Gens.

Das resultierende Plasmid wurde mit pCB289-13 bezeichnet. Trotz der in E. coli zu erwartenden Expression des Enzyms I-PpoI wurden keine Beeinträchtigungen des Wachstums festgestellt. Von der 35 KpnI bis zur SacI Schnittstelle ergab sich durch SEQ ID NO: 4 beschriebene Sequenz (Vektorrückgrat ist unverändert das des pBluescript KS).

5.2 Erstellen eines Transformationsvektors für artifizielles 40 Homing mit beidseitig die Insertionssequenz flankierende homologe Regionen

#### I) ohne I-PpoI in der Insertionssequenz

45 Bereiche um die I-PpoI Erkennungsregion aus pCB199-3 wurden mit PstI und SacI ausgeschnitten und in die PstI und SacI Schnittstellen des pBluescript legiert. Anschließend wurden aus diesem

## 96

Vektor nicht benötigte Schnittstellen entfernt, indem er mit PstI und Bsp120I linearisiert wurde und nach einer Behandlung mit dem Klenow-Fragment der Vektor wieder rezirkularisiert wurde. Mit Hilfe kommerziell verfügbarem Enzym I-PpoI (PROMEGA GmbH, 5 Mannheim, Deutschland) wurde in dem erhaltenen Vektor die entsprechende Erkennungsregion geschnitten und dort hinein mittels der synthetischen Oligos p190 und p191 weitere Schnittstellen eingebracht.

10 Oligo p190: 5'-GTCGACAGATCTTTAA-3' (SEQ ID NO: 20)

Oligo p191: 5'-AGATCTGTCGACTTAA-3' (SEQ ID NO: 21)

In die so eingebrachten Schnittstellen SalI und BglII wurde eine 15 Expressionskassette als BglII / XhoI Fragment (SEQ ID NO: 6) bestehend aus folgenden Elementen eingebracht:

- a) Prps16 Promotor (komplementär 1033-1139)
- 20 b) 5'rbcL (komplementär 1007-1024) mit 18 bp codierend für die 6 AS (komplementär 989-1006)
- c) nptII Gen (komplementär 185-988)
- 25 d) 3'rbcL (komplementär 6-143)

Der resultierende Vektor wurde mit pCB304-25 bezeichnet und verlieh auch E. coli Zellen Kanamycin Resistenz. Dieser Vektor wird durch kommerziell verfügbares I-PpoI nicht mehr linearisiert. Das 30 gesamte Insert des Vektors pCB304-25 (Grundgerüst pBluescript; von SacI bis KpnI entsprechend die MCS ersetzend) ist durch SEQ ID NO: 62 beschrieben und umfasst somit folgende Elemente:

- a) Position bp 19 bis 110: 3'psbA aus Tabak
- 35 b) Position bp 149 bis 160: nicht funktionelle "Halbseite" der I-PpoI Erkennungssequenz
- c) Position bp 171 bis 277: Prps16 Promotor
- 40 d) Position bp 286 bis 303: 5'rbcL Sequenz gefolgt von 18bp kodierend für die ersten 6 Amino säuren des von rbcL Proteins (bp 304-321)
- 45 e) Position bp 322 bis 1125: nptII

97

- f) Position bp 1167 bis 1304: 3'rbcl  
g) Position bp 1310 bis 1319: nicht funktionelle "Halbseite"  
der I-PpoI Erkennungssequenz

5

II) Mit I-PpoI in der Insertionssequenz.

Mit Hilfe der BamHI und EcoRI Schnittstellen wurde in den Vektor pCB304-25 zusätzlich ein BglII / MunI Fragment eingebettet, welches für eine 5'psbA - I-PpoI Fusion kodierte.  
Der entstandene Vektor pCB320-192 exprimierte unter Kontrolle des Prps16 Promotor also das nptII Gen und die I-PpoI Homing Endonuklease. Das Bgl II / Mun I Fragment ist durch SEQ ID NO: 63 wiedergegeben und umfasst folgende Elemente:

15

- a) Position bp 6 bis 82: 5'psbA  
b) Position bp 83 bis 574: I-PpoI

20 5.3 Erstellen eines Transformationsvektors für artifizielles Homing mit einer einseitig die Insertionssequenz flankierenden homologen Region

Aus dem Vektor pCB320-192 wurden durch Restriktion mit KpnI und BglII die stromaufwärts des Prps16 Promotor gelegenen Elemente, die homolog zu denen in den Masterpflanzen CB199NTH sind, entfernt. An deren Stelle wurde mittels synthetischer Oligonukleotide p199 und p200 eine BstXI Schnittstelle dort eingebracht.

30 p199 5'-GATCTCCAGTTAACCTGGGGTAC-3' (SEQ ID NO: 22)

p200 5'-CCCAGTTAACCTGGA-3' (SEQ ID NO: 23)

Durch Schneiden mit BstXI kann man nun DNA-Enden erzeugen, die kompatibel zu denen sind, die durch Restriktion mit I-PpoI entstehen. Der erhaltene Vektor wurde mit pCB322-1 bezeichnet. Beispielsweise mit den Enzymen BstXI und SacI kann aus diesem Vektor ein Fragment gewonnen werden, welches an der einen Seite ein zu DNA, die mit I-PpoI an seiner Core-Erkennungsregion geschnitten wurde, kompatibles Ende und an der anderen Seite Homologie zu Plastomsequenzen der Masterpflanzen CB199NTH besitzt.

5.4 Erstellen eines Transformationsvektors für artificielles Homing ohne homologe Bereiche um die Insertionssequenz herum

Aus dem Vektor pCB322-1 wurde mit SacI und BspTI der verbliebende Teil, der homolog zu rekombinanten Plastidensequenzen der Masterpflanzen CB199NTH ist, entfernt. Gleichzeitig wurde durch einbringen synthetischer Oligonukleotide p218 und p219 hier eine BstXI Schnittstelle erzeugt, die nach Schneiden mit BstXI DNA-Enden generiert, welche kompatibel zu mit I-PpoI geschnittener DNA sind. Der resultierende Vektor wurde mit pCB347-33 gekennzeichnet.

p218 5'-TTAAGCCAGTTAACTGGGCGGAGCT-3' (SEQ ID NO: 24)

15 p219 5'-CCGCCAGTTAACTGGC-3' (SEQ ID NO: 25)

Mit dem Enzym BstXI kann aus diesem Vektor ein Fragment isoliert werden, welches auf beiden Seiten DNA-Enden aufweist, welche kompatibel zu den DNA-Enden sind, die das I-PpoI Enzym an seiner Core-Erkennungsregion erzeugt.

Beispiel 6: Ausnutzen der Masterpflanzen CB199NTH für eine Plastidentransformation mittels DSB-Induktion

25 6.1 Nutzen der unter 5.1 und 5.2 erstellten Vektoren zur Transformation der Masterpflanzen CB199NTH durch Ausnutzen des DSBI-Enzyms I-PpoI

Das Plasmid pCB304-24 wurde gleichzeitig mit dem Plasmid 30 pCB289-13 wie unter Beispiel 4 beschrieben auf Goldpartikel aufgebracht und in entsprechend den Ausführungen in Beispiel 4 behandelte Explantate der Masterpflanzen CB199NTH geschossen. In Abweichung zu den Beschreibungen in Beispiel 4 erfolgt zunächst eine Inkubation für 10 Tage auf dem Regenerationsmedium ohne 35 Antibiotikum, später wird Kanamycin in einer Konzentration von 50 mg/L verwendet (im Gegensatz zu den in Beispiel 4 angegebenen 500 mg/L Spectinomycin).

Das Plasmid pCB320-192 wurde wie unter Beispiel 4 beschrieben auf 40 Goldpartikel aufgebracht. Nach dem Waschen mit Ethanol wurden noch 20 U kommerziell verfügbarem I-PpoI Enzym (Promega) hinzugegeben. Die weitere Behandlung erfolgt wie oben beschrieben.

Ferner wurde in einem anderen Ansatz 0,5 µg eines mit Hilfe 45 der T7-Polymerase in vitro erstellten Transkriptes gleichzeitig mit dem Plasmid pCB320-192 auf die Goldpartikel aufgebracht. Als Matrize für die in vitro Transkription diente mit HindIII

99

linearisierte DNA des Plasmides pCB289-13. Das so erstellte Transkript codiert also für I-PpoI. Nach dem Beschuss werden die Explantate der Masterpflanzen wie oben beschrieben weiter behandelt.

5

6.2 Nutzen der unter 5.1 und 5.3 erstellten Vektoren zur Transformation der Masterpflanzen CB199NTH durch Ausnutzen des DSBI-Enzyms I-PpoI und nur einseitig vorhandener homologer Regionen

10

Ein mit BstXI und SacI aus dem Plasmid pCB322-1 ausgeschnittenes Fragment wurde aus einem Agarosegel eluiert. Dieses Fragment wurde anschließend gleichzeitig mit 1 µg in vitro Transkript von mit HindIII linearisierten Plasmid pCB289-13 (vgl. Beispiel 6.1) 15 auf Goldpartikel gebracht. Nach dem Beschuss der Explantate der Masterpflanzen CB199NTH wird die weitere Behandlung wie unter Beispiel 6.1 beschrieben vorgenommen.

20

6.3 Nutzen der unter 5.1 und 5.4 erstellten Vektoren zur Transformation der Masterpflanzen CB199NTH durch Ausnutzen des DSBI-Enzyms I-PpoI ohne homologe Regionen

Aus dem Plasmid pCB347-33 wurde mittels BstXI die Insertionssequenz ausgeschnitten und aus einem Agarosegel eluiert. Dieses 25 Fragment wurde gleichzeitig mit 1 µg in vitro Transkript des mit HindIII linearisierten Plasmides pCB289-13 auf Goldpartikel aufgebracht. Beschuss und nachfolgende Behandlung erfolgt wie unter Beispiel 6.1 ausgeführt.

30 Beispiel 7: Identifizieren von natürlich vorkommenden, endogenen Erkennungsregionen für Homing Endonukleasen in Plastomen verschiedener Pflanzenspezies

Obwohl in den Plastiden höherer Pflanzen keine Homing Endonukleasen bekannt sind, wurden bekannte Plastomsequenzen auf das Vorhandensein von Erkennungsregionen für Homing Endonukleasen hin untersucht. Dies erfolgte mit Hilfe des Computerprogramms SeqMan II (DNASTAR Inc.). Die Erkennungsregionen, die auf diese Weise identifiziert wurden, sind in Tabelle 1 zusammengefasst.

40

Es war aufgrund der Computeranalyse nicht möglich, zu entscheiden, ob I-SceI im Plastidengenom eine Erkennungsregion besitzt oder nicht. Die Region, die mit höchster Wahrscheinlichkeit als Erkennungsregion fungieren kann, wurde synthetisch erstellt und als Oligonukleotide p276 und p277 in die XbaI und XhoI Schnittstelle des pBluescript integriert. Das resultierende Plasmid pCB414-1 wurde anschließend mit Hilfe kommerziell ver-

## 100

fügbarem Enzym I-SceI (Roche) auf das Vorhandensein einer funktionellen Schnittstelle hin analysiert. Tatsächlich wurde das Plasmid durch I-SceI linearisiert. Daraus kann gefolgt werden, dass in Plastiden exprimiertes I-SceI ebenfalls diese Sequenz erkennt und schneidet. Damit ist eine weitere endogene DSB-Erkennungssequenz für ein DSBI-Enzym identifiziert worden.

5 p276 5'-TCGAGAAGATCAGCCTGTTATCCCTAGAGTAAC-3' (SEQ ID NO: 26)

10 p277 5'-CTAGAGTTACTCTAGGGATAAACAGGCTGATCTTC-3' (SEQ ID NO: 27)

Beispiel 8: Klonieren von homologen Bereichen aus dem Tabakplastom, die die endogene Erkennungsregion für die Homing Endonuklease I-CpaI flankieren.

15

Stromauf- und stromabwärts der I-CpaI Erkennungsregion wurden DNA-Fragmente aus der 23S rDNA des Tabakplastom mittels PCR unter Verwendung der Primer p93 und p97 bzw. p98 und p95 amplifiziert.

20 p93: AAAGATCTCCTCACAAAGGGGTGCG (SEQ ID NO: 64)

p97: TCGAAGACTTAGGACCCTTATAG (SEQ ID NO: 65)

p98: AGGAAGACCTTGTCGGGTAAGTTCCG (SEQ ID NO: 66)

25

p95: CTCAATTGGGGTCTCTGTCCAGGTGCAGG (SEQ ID NO: 67)

Die so erhaltenen Fragmente wurde zur Konstruktion des Vektors pCB270-1 verwendet. Angegeben ist das Fragment von BssHII bis 30 BssHII des pBlueScript Vektors (SEQ ID NO: 8), wobei das angegebene 5' Ende der Sequenz an der BssHII Schnittstelle zu finden ist, die näher am 3' Ende des lacZ Genes liegt.

Zwischen den stromaup- und stromabwärts der I-CpaI Erkennungs- 35 regionen gelegenen DNA Fragmenten wurden zwei BpiI Schnittstellen eingeführt. BpiI erzeugt überhängende Enden, die außerhalb ihrer Erkennungsregion liegen. Durch dieses Vorgehen konnte gewährleistet werden, dass der Vektor pCB270-1 für die nachfolgende Integration verschiedener Introns gleichermaßen genutzt werden 40 kann. Dazu werden an den zu klonierenden Introns einfach überhängende Enden erzeugt, die kompatibel zu den durch BpiI erzeugten Enden im Vektor pCB270-1 sind. Ferner addiert man an die Introns die entsprechenden Nukleotide, die zwischen den beiden Fragmenten der 23S rDNA im Vektor pCB270-1 fehlen. Die ausgewählten Bereiche sind so hoch konserviert, dass aus anderen Pflanzen- 45 spezies keine neuen Regionen amplifiziert werden brauchen. Ferner ist in die stromabwärts der I-CpaI Erkennungsregion gelegene

## 101

Sequenz über die PCR Strategie eine Punktmutation in die 23S rDNA Fragment eingebaut worden, wie sie auch in Lincomycin resistenten Mutanten gefunden wurde. Die in den pBluescript Vektor inserierte Sequenz des Vektors pCB270-1 ist unter der SEQ ID NO: 8 5 wiedergegeben. Die Sequenz umfasst nachfolgende Elemente:

- Fragment der 23SrDNA stromaufwärts der I-CpaI Erkennungsregion: Nukleotide 37 bis 194
- 10 - Fragment der 23SrDNA stromaufwärts der I-CpaI Erkennungsregion: Nukleotide 237 bis 359
- Punktmutation für Lincomycinresistenz: 352 (nativ ist hier A),

15 Vektor pCB234-1 ist genauso aufgebaut wie der Vektor pCB270-1, er besitzt lediglich stromabwärts der unten angegeben Sequenz noch je eine Erkennungsregion für die Restriktionsendonukleasen XbaI und SacI.

20 Beispiel 9: Klonieren des CpLSU2 Intron einschließlich der Homing Endonuklease I-CpaI

Aus DNA der Alge Chlamydomonas pallidostigmatica (Sammlung von Algenkulturen Göttingen, Culture Collection of Algae at the 25 University Göttingen, SAG Nummer 9.83, Chlamydomonas segnis, authentic strain of Chlamydomonas pallidostigmatica King) wurde mittels PCR mit den Oligonukleotiden p91 und p92 das CpLSU2 Intron amplifiziert (SEQ ID NO: 15).

30 p91 5'-AGAAGACGATCCTAAGG-3' (SEQ ID NO: 28)

p92 5'-TGAAGACTTGACAAGGAATTTCGC-3' (SEQ ID NO: 29)

Die Wahl der Oligonukleotide war derart, dass wie oben 35 beschrieben eine Klonierung in die BpiI Schnittstellen des Vektors pCB234-1 möglich war. Die Sequenz umfasst nachfolgende Elemente:

- Position 9-17 - Anteil der Tabak 23S rDNA, der in pCB234-1 40 fehlt
- CpLSU2-Intron: Position 18-893
- I-CpaI ORF: Position 377-835

102

- Position 894-909 - Anteil der Tabak 23S rDNA, der in pCB234-1 fehlt

Dieses Fragment, welches das CplSU2 Intron umfasst, wurde in 5 das Rückgrat des Vektors pGEMTeasy (Promega) kloniert (Vektor PCB141-3). Das gesamte Fragment wurde mit BpiI aus diesem Vektor ausgeschnitten und in den mit BpiI linearisierten Vektor pCB234-1 kloniert. Der resultierende Vektor wurde mit mit pCB254-2 bezeichnet.

10

Beispiel 10: Nukleäres LSU-rRNA Intron aus *Tetrahymena thermophila*

10.1 Klonieren des nukleären LSU-rRNA Intron aus *Tetrahymena thermophila*

15

Mittels PCR wurde aus dem Organismus *Tetrahymena thermophila* das LSU-rRNA Intron amplifiziert. Die Oligonukleotide p102 und p103 wurden wieder derart gewählt, dass an das zu amplifizierende Intron die in pCB234-1 fehlenden Nukleotide der Tabak 23S rDNA 20 angefügt wurden.

p102 (SEQ ID NO: 30):

5'-AGAAGACGATCCTAAATAGCAATATTACCTTGGGACCAAAAGTTATCAGGCATG-3'

25 p103 (SEQ ID NO: 31):

5'TGAAGACTTGACAAGGAATTCGTACCCTTGAGTACTCCAAAACTAATC-3'

Außerdem ist die *internal guide sequence* (Unterstrichen in p102) durch die Wahl des Oligonukleotides p102 derart gegenüber 30 dem Wildtyp mutiert, dass ein Spleißen dieses Introns an der gewünschten Position in der 23SrDNA vom Tabak möglich ist. Die unter SEQ ID NO: 7 angegebene Sequenz - angegeben ist das PCR Fragment von BpiI bis BpiI Schnittstelle - wurde in das Rückgrat des Vektors pGEMTeasy kloniert. Der erhaltene Vektor wurde mit 35 pCB220-17 bezeichnet. Die Sequenz umfasst nachfolgende Elemente:

- Position 9-12 - Anteil der Tabak 23S rDNA, der in pCB234-1 fehlt

40 - LSU-Intron: Position 13-425

- Position 426-446 - Anteil der Tabak 23S rDNA, der in pCB234-1 fehlt

45 Das *Tetrahymena* LSU-Intron wurde einschließlich der angefügten, flankierenden Sequenzen aus dem Vektor pCB220-17 mit BpiI ausgeschnitten und in die BpiI Schnittstellen des Vektors pCB234-1

103

insertiert. Das entstandene Produkt wurde mit pCB255-1 bezeichnet.

10.2 Indirekter Nachweis der Spleißaktivität des Tetrahymena  
5           LSU Intron in E. coli

Zum indirekten Nachweis, dass das modifizierte Intron in der vorbestimmten Umgebung innerhalb der I-CpaI Schnittstelle tatsächlich Spleißen kann, wurde das modifizierte Tetrahymena 10 Intron aus pCB220-17 zusammen mit einem Anteil, der die I-CpaI Erkennungsregion aus der Tabak 23SrDNA umgibt, so in das lacZ Gen des pBluescript kloniert, dass im Fall des Spleißens dieses 15 Introns in E. coli (Stamm XL1-Blue) ein funktionelles lacZ Peptid gebildet werden kann. Dessen Expression kann in geeigneten 20 Stämmen durch dem Fachmann geläufige Verfahren durch Umsetzen der Substanz 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranosid (X-Gal) im Medium zu einem blauen Farbstoff nachgewiesen werden. Der entsprechende Vektor wurde mit pCB315-1 bezeichnet. Das lacZ Gen einschließlich des Introns im Vektor pCB315-1 ist durch 25 SEQ ID NO: 9 beschrieben. Das Vektorrückgrat ist identisch zum pBluescript. Die Sequenz umfasst nachfolgende Elemente:

- lacZ-5' Anteil: komplementär (789-765)
- 25 - multiple Klonierungsstelle aus dem pBluescript: komplementär (764-692)
- 23SRDNA Fragment stromaufwärts und einschließlich der I-CpaI Erkennungsregion: komplementär (691-682)
- 30 - modifiziertes Tetrahymena Intron komplementär (681-269)
- 23SRDNA Fragment stromabwärts und einschließlich der I-CpaI Erkennungsregion: komplementär (268-244)
- 35 - multiple Klonierungsstelle aus dem pBluescript: komplementär (243-168)
- lacZ-5' Anteil: komplementär (167-1)

40 Zur Kontrolle wurde ein Plasmid erstellt, das dem pCB315-1 entspricht, nur dass in diesem Plasmid (pCB305-1) das Element für das modifizierte Tetrahymena Intron fehlte. pCB305-1 diente somit als Positivkontrolle, um zu zeigen, dass das lacZ mit den im 45 Leseraster eingebauten Nukleotiden aus der 23SrDNA vom Tabakplastom noch funktionell ist. Dies spiegelt die Situation nach einem korrekten Spleißen des Tetrahymena Intron wieder. XL1-Blue

## 104

kompetente Zellen wurden mittels einer dem Fachmann geläufigen Methode mit den Plasmiden pCB315-1 und pCB305-1 transformiert. Je eine Einzelkolonie wurde auf LB (Bactotrypton: 10 g/L, Hefeextrakt: 5 g/L, NaCl: 10 g/L, pH7,5) Platten, die 15g/L Bacto-  
5 Agar, 40 µg/L Ampicillin, 75 µg/L IPTG (Isopropyl-β-D-Thiogalactopyranosid) und 80 µg/L X-Gal enthielten, über Nacht bei 37°C inkubiert. Tatsächlich färbten beide Klone blau, was nahe legt,  
dass das modifizierte Tetrahymena Intron in der nicht natürlichen Umgebung der Tabak 23SrDNA im heterologen Organismus E. coli  
10 gespleißt wurde.

10.3: Einfügen von weiteren Sequenzen in das Tetrahymena LSU  
Intron

15 Zusätzlich zu den Versuchen in Beispiel 10.2 wurde untersucht, ob weitere Elemente in das modifizierte Tetrahymena Intron eingebaut werden können, ohne die Spleißaktivität zu zerstören. Dazu wurde pCB315-1 mit BglII linearisiert und die überhängenden Enden mit Hilfe des Klenow-Fragmentes aufgefüllt. In diesen Vektor wurde  
20 so dann ein XhoI-SacI Fragment, wie es in pCB199-3 vorkommt, ebenfalls nach Behandlung mit dem Klenow-Fragment kloniert. Durch diese Klonierung wurde also ein 229bp großes Fragment in das modifizierte Intron eingesetzt. Dieses Fragment enthält eine I-PpoI Erkennungsregion. Unabhängig von der Orientierung mit der  
25 das 229bp Fragment in das Tetrahymena Intron insertierte, konnte in dem Test wie unter Beispiel 10.2 beschrieben eine Blaufärbung nachgewiesen werden. Dies legt nahe, dass das Tetrahymena-Intron sowohl an der gewünschten Region in der 23S rDNA Spleißen kann, als auch zusätzliche genetische Information aufnehmen kann und  
30 trotzdem weiterhin eine Spleißaktivität gegeben ist.

10.4 Transformation einer natürlichen Masterpflanze und  
Zerstörung der endogenen I-CpaI Erkennungsregion mit  
dem modifizierten Tetrahymena Intron

35 Nach dem in Beispiel 8 beschriebenen Prinzip wurde das modifizierte Tetrahymena Intron aus dem Vektor pCB220-17 mit BpiI ausgeschnitten und in den mit BpiI linearisierten Vektor pCB234-1 kloniert. Der resultierende Vektor wurde pCB255-1 genannt.  
40 pCB255-1 wird gleichzeitig mit in vitro Transkript vom pCB262-5 (mit SalI linearisiert, Nutzen der T7 Polymerase) nach dem in Beispiel 4 beschriebenen Verfahren auf Goldpartikel aufgebracht. Anschließend werden diese Goldpartikel analog zum Verfahren beschrieben in Beispiel 4 in Tabak Pflanzen, Varietät Petit Havana,  
45 geschossen. Gegebenenfalls können die Explantate auf Lincomycin (250 bis 500 mg/L) selektiert werden.

105

Beispiel 11: Ll.LtrB Intron aus *Lactococcus lactis*

Das Ll.LtrB Intron wurde einschließlich weniger Basen der flankierenden Exonsequenzen aus *Lactococcus lactis* mittels PCR durch Ausnutzen der Primer p207 und p208 amplifiziert. Das PCR Produkt wurde in den Vektor pCR2.1TA (Invitrogen) kloniert (pCB345-34) und sequenziert (SEQ ID NO: 10). Es wurden wenige Abweichungen im Vergleich zur veröffentlichten Sequenz gefunden.

10 p207 5'-GAGAAGACATTCTAACACATCCATAACGTGCG-3' (SEQ ID NO: 32)

p208 5'-TGAAGACTTGACATTTGATATGGTGAAGTAGG-3' (SEQ ID NO: 33)

Das klonierte Fragment in pCB345-34 (von der EcoRI bis zur EcoRI 15 Schnittstelle des pCR2.1TA Vektors) ist durch SEQ ID NO: 10 wiedergegeben. Der Rest des Vektors ist identisch zum Rückgrat des pCR2.1TA. Die Sequenz umfasst nachfolgende Elemente:

- Anteil des natürlichen 5'Exon: komplementär (2540-2527)
- 20 - Intron Ll.LtrB: komplementär (2526-35)
- ORF im Intron: komplementär (1953-154)
- 25 - Anteil des natürlichen 3'Exon: komplementär (34-28)

Beispiel 12: Erstellen eines weiteren Derivates des Tetrahymena LSU Intron und Einbau eines Fremdgens in dieses Intron-Derivat, sowie Transformation in natürliche 30 Masterpflanzen

Gemäß einer bevorzugten Ausführungsform dieser Erfindung wurde ein artifizielles Intron erstellt, welches genau an der Stelle in das Plastidengenom eingebaut werden kann, an der sich auch das 35 natürliche Intron zugehörig zu der betrachteten DSB-Erkennungssequenz befindet. Im vorliegenden Beispiel wurde das Tetrahymena LSU Intron derart modifiziert, dass es an der mit „^“ gekennzeichneten Position in der im Rahmen dieser Erfindung identifizierten Erkennungsstelle für das DSBI-Enzym I-CpaI im Plastiden-40 genom höherer Pflanzen spleißen kann: CGATCCTAACGGT^AGCGAAATTCA.

Anschließend wurde das Gen codierend für I-CpaI einschließlich einer RBS in das Intron eingebaut. So wurde ein spleißaktives, ein Fremdgen tragendes Intron erstellt, welches mittels des im 45 Rahmen dieser Erfindung gefundenen Prozesses in die Plastiden einer natürlichen Masterpflanze inmitten eines essentiellen Gens (codierend für die 23S rRNA) eingebaut werden kann.

## 106

## 12.1: Erstellen eines weiteren Tetrahymena LSU Intron-Derivates

Um ein an einer vordefinierten Insertionsstelle ein funktionelles Intronderivat zu erhalten, muß die *internal guide sequence* (IGS) 5 derart angepasst werden, dass sie Basenpaarungen mit dem 5' und 3' Exon eingehen kann. Fig. 10 illustriert, wie diese Anpassung im vorliegenden Beispiel erfolgte, um ein Tetrahymena Intron-Derivat zu erstellen, welches an der natürlichen Insertionsstelle des CpLSU5 Intron innerhalb der I-CpaI Erkennungsregion spleissen 10 kann. Analog kann eine Anpassung an jede beliebige Insertionsstelle vorgenommen werden. Das im Rahmen dieses Beispiels erstellte Intron wurde mit TetIVS2a bezeichnet und ist durch SEQ ID NO: 73 beschrieben.

## 15 12.2: Indirekter Nachweis der Spleißaktivität des TetIVS2a Introns in E. coli

Analog zum Beispiel 10.2 wurde das TetIVS2a in das lacZ Gen des pBluescript eingebaut. Nach entsprechender Inkubation von E. coli 20 XL1-blue Zellen, die das Plasmid pCB459-1 enthielten, zeigte eine Blaufärbung die Spleißaktivität des TetIVS2a Intron an der gewünschten Position an.

Bestandteile des Inserts aus Plasmides pCB459-1 (SEQ ID NO: 74; 25 Rückgrat entspricht pBluescript)

- lacZ-3' Anteil einschließlich Teile der multiplen Klonierungsstelle aus dem pBluescript (komplementär Pos. 1-254)
- Sequenz aus der I-CpaI Erkennungsregion (komplementär Pos. 30 254-265)
- TetIVS2a (komplementär Pos. 266-678)
- Sequenz aus der I-CpaI Erkennungsregion (komplementär Pos. 679-687)
- lacZ-5' Anteil einschließlich Teile der multiplen Klonierungsstelle aus dem pBluescript (komplementär Pos. 688-791)

## 35 12.3: Einführen einer weiterer genetischer Information in das TetIVS2a Intron und Nachweis der Spleißaktivität in E. coli

40 Im Rahmen dieses Beispiels wird das Gen kodierend für das DSBI-Enzym I-CpaI in das TetIVS2a eingebracht, ohne dass dieses seine Spleißaktivität an besagter Position innerhalb der I-CpaI Erkennungsregion verliert. Dazu wurde zunächst in den Sequenzabschnitt 45 des TetIVS2a, welcher dem Loop L8 im Tetrahymena LSU Intron entspricht, eine BclI Schnittstelle mittels PCR eingefügt. Anschließend wurde in diese BclI Schnittstelle ein nicht funktionelles

107

Derivat des Gens codierend für I-CpaI eingebaut. Da die Expression von I-CpaI in *E. coli* toxisch ist, mußte für den Spleißtest in *E. coli* ein nicht-funktionelles Gen genutzt werden, welches zuvor dadurch erzeugt wurde, daß das entsprechende Gen an der EcoRI 5 Schnittstelle linearisiert wurde, die überstehenden Enden durch Klenow-Behandlung geglättet wurden und anschließend die Genabschnitte wieder zusammen ligiert wurden. Dadurch wurde ein Rasterschub im Leserahmen des Gens erzeugt. Durch den Einbau des besagten Introns mit dem nicht-funktionellen Gen in das lacZ Gen 10 des pBluescript erhielten wir das Plasmid pCB478-3 und konnten analog zu Beispiel 12.2 mittels einer Blaufärbung entsprechender Kolonien die Spleißaktivität auch dieses Introns in *E. coli* an der gewünschten Position innerhalb der I-CpaI Erkennungsstelle nachweisen. Da sich das funktionelle Gen kodierend für I-CpaI nur 15 durch 4 Basen von dem in pCB478-3 genutzten nicht-funktionellen Gen unterscheidet, kann davon ausgegangen werden, daß das Intron auch nach Einbau des funktionellen I-CpaI Gens anstelle des nicht-funktionellen weiterhin die gewünschte Spleißaktivität besitzt.

20

Bestandteile des Inserts aus Plasmides pCB478-3 (SEQ ID NO: 75; Rückgrat entspricht pBluescript)

- lacZ-3' Anteil einschließlich Teile der multiplen Klonierungsstelle aus dem pBluescript (komplementär Pos. 1-265)
- Sequenz aus der I-CpaI Erkennungsregion (komplementär Pos. 256-265)
- TetIVS2a (komplementär Pos. 266-1178), darin enthalten ein nicht-funktionelles Gen für I-CpaI (komplementär Pos. 399-861) sowie einer RBS stromaufwärts des nicht-funktionellen I-CpaI Gens (komplementär 866-870)
- Sequenz aus der I-CpaI Erkennungsregion (komplementär Pos. 1179-1187)
- lacZ-5' Anteil einschließlich Teile der multiplen Klonierungsstelle aus dem pBluescript (komplementär Pos. 1179-1291)

#### 12.4: Transformation eines sich selbst verbreitenden, artifiziellen Intron in einer natürlichen Masterpflanzen

40 Nachdem gezeigt werden konnte, dass das TetIVS2a Intron an der gewünschten Position spleißen kann und ohne diese Spleißaktivität zu verlieren gleichzeitig weitere genetische Information aufnehmen kann, wurde ein Konstrukt erstellt, welches es ermöglichen soll, das I-CpaI Gen als ein Fremdgen mittels des im Rahmen die- 45 ser Erfindung beschriebenen Verfahrens in das Plastidengenom einzubauen, ohne einen Selektionsmarker zu nutzen. Dazu wurde zu- nächst der Vektor pCB492-25 erstellt, welcher ein Insert mit fol-

108

gende Elemente umfasst (SEQ ID NO: 76; Rückgrat entspricht dem des pBluescript; Sequenz ist angegeben von BssHII bis BssHII im pBluescript, wobei die hier am 5' Ende angegebene BssHII Schnittstelle, diejenige im pBluescript ist, die näher am 3' Ende des 5 lacZ Gens lokalisiert ist):

- 23SrDNA Fragment stromaufwärts und einschließlich der I-CpaI Erkennungsregion (Pos. 37-203)
- TetIVS2a (Pos. 204-1112) mit eingefügtem Gen codierend für I-CpaI (Pos. 521-979) sowie RBS (Pos. 512-516)
- 10 - 23SrDNA Fragment stromabwärts und einschließlich der I-CpaI Erkennungsregion (Pos. 1113-1247)

Um eine Expression des I-CpaI direkt nach dem Einbringen in Pla-  
15 stiden von natürlichen Masterpflanzen zu gewährleisten, wurde in vitro ein Promotor stromaufwärts des besagten Intron-Cpa Kon-  
struktes mittels PCR angefügt. Dazu wurden die Primer p396 und  
p95 sowie pCB492-25 als Matritze genutzt.

20 p396 (SEQ ID NO: 77):

5'-TAGTAAATGACAATTTCCTCTGAATTATATAATTACATGGCGACTGTTACCAAAAC-3

p95 (SEQ ID NO: 78): 5'-CTCAATTGGGTCTCTGTCCAGGTGCAGG-3'

25 Das entstandene PCR Produkt erhielt die Bezeichnung Prom-TetIVS2a-Cpa, ist beschrieben durch SEQ ID NO: 79 und umfasste fol-  
gende Elemente:

- synthetischer Promotor (Pos. 8-40)
- 30 - 23SrDNA aus Tabak stromaufwärts und einschließlich Anteile der I-CpaI Erkennungsregion (Pos. 41-207)
- TetIVS2a (Pos. 208-1116) umfassend Gen codierend für I-CpaI (Pos. 525-983) sowie RBS (Pos. 516-520)
- 23SrDNA aus Tabak stromabwärts und einschließlich Anteile der I-CpaI Erkennungsregion (Pos. 1117-1243)

Das Plasmid pCB492-25 wurde gleichzeitig mit dem beschriebenen PCR-Produkt Prom-TetIVS2a-Cpa wie in Beispiel 4 beschrieben auf Goldpartikel aufgebracht und anschließend auf Wildtyp Tabak ge-  
40 schossen. Durch die Expression des I-CpaI Enzyms soll ein Doppel-  
strangbruch in der 23SrDNA erfolgen, welcher durch das einge-  
brachte PCR Produkt oder durch die eingebrachte Insertionssequenz des Plasmides pCB492-25 repariert wird. Dadurch wird die natürli-  
cherweise vorhandene I-CpaI Erkennungsregion in den bereits  
45 transformierten Plastomkopien inaktiviert. Ohne jedlichen Selektionsdruck wurden Pflanzen regeneriert und diese werden mittels

109

PCR auf die Anwesenheit der Insertionssequenz im Plastom getestet.

5 Beispiel 13: Verbreitung des modifizierten Tetrahymena LSU Intron aus pCB255-1 in einer natürlichen Masterpflanze durch Expression des DSBI-Enzyms I-CpaI in trans

10 Im Rahmen dieses Beispiels wird gezeigt, wie ein DSBI-Enzym in den Plastiden von Masterpflanzen zur Expression gebracht werden kann, um eine Insertionssequenz in den Kopien der Masterpflanze 15 effizient zu verbreiten.

13.1: Erstellen eines Vektors für die Plastidentransformation, der die Expression der Homing Endonuklease I-CpaI in Pla-  
15 stiden erlaubt

Zunächst wurde ein Vektor erstellt, der für den Selektionsmarker aadA sowie das DSBI-Enzym I-CpaI kodiert. Da die Expression des I-CpaI Enzyms in E. coli letal ist, wurde der accD Promotor ge-  
20 wählt, um die Expression dieses Enzyms in den Plastiden zu ermöglichen, aber die Expression in E. coli zu verhindern. So konnte dieser Vektor auf konventionelle Art und Weise mit E. coli als Wirtsorganismus erstellt und amplifiziert werden. Der resultierende Vektor wurde mit pCB435-45 bezeichnet und umfaßte ein In-  
25 sert gemäß SEQ ID NO: 80 mit folgenden Elementen:

- Rechte Zielregion (wie in pCB42-94, siehe oben; komplementär Pos. 66-1403)
- Promotor PaccD (Pos. 1422-1478)
- 30 - RBS (Pos. 1500-1504)
- Gen codierend für I-CpaI (Pos. 1509-1967)
- Expressionskassette für das Marker-Gen aadA bestehend aus:
  - der 3' Bereich des psbA Gens (komplementär Pos. 2065-1974)
  - aadA Gen (komplementär Pos. 2872-2078)
- 35 - 5' untranslatierte Bereiche des Tabak rbcL Gens (komplementär Pos. 2890-2873), teilweise mutiert
- Promotor des Gens für die 16SrRNA (komplementär Pos. 2987 bis 2897)
- Linke Zielregion (wie in pCB42-94, siehe oben; komplementär Pos. 3863-3007)
- 40 - Teile des pBluescript (einschließlich „origin of replication“; Pos. 3864-4746 und 1-65)

110

## 13.2 Co-Transformation von pCB435-45 und pCB255-1 in natürliche Masterpflanzen

Entsprechend den Ausführungen in Beispiel 4 wurden die Plasmide  
5 pCB435-45 und pCB255-1 gleichzeitig auf Goldpartikel aufgebracht  
und so dann mittels der Partikelkanone in Plastiden von Tabak-  
blättern eingebracht. Analog zu Beispiel 4 wurden Transplastome  
Pflanzen auf Regenerationsmedium zuzüglich 500 mg/L Spectinomycin  
selektiert. Sobald Pflänzchen entstanden waren, wurden diese auf  
10 Anzuchtmittel zuzüglich 500 mg/L Spectinomycin umgesetzt und  
Blattmaterial geerntet. Dieses Blattmaterial wurde mittels Sou-  
thernanalyse unter Verwendung des Dig-Easy Hyb® (Roche Diagno-  
stics; Mannheim) in Hinblick auf den Einbau der beiden Plasmide  
in das Plastidengenom analysiert. Um den Anteil der Plastomkopien  
15 zu bestimmen, die bzgl. der Insertionssequenz aus pCB435-45  
transgen waren, wurde eine Sonde benutzt, die eine Sequenz gemäß  
SEQ ID NO: 81 besaß (gerichtet gegen Teile der 16SrDNA).

Um den Anteil der Plastomkopien zu bestimmen, die bzgl. der In-  
20 sertionssequenz aus pCB255-1 transgen waren, wurde eine Sonde be-  
nutzt, die eine Sequenz gemäß SEQ ID NO: 82 besaß (gerichtet ge-  
gen Teile der 23SrDNA).

Fig. 11 zeigt, dass in diesem Experiment 2 Linien  
25 (CB255+435NTH-19 und -20) identifiziert wurden, die sowohl bzgl.  
der Insertionssequenz aus pCB435-45 als auch der aus pCB255-1  
transgen sind. Weiter konnte so gezeigt werden, dass überraschend-  
weise die Insertionssequenz aus pCB255-1 (modifiziertes Tetra-  
hymena LSU Intron) bereits in mehr Kopien des Plastidengenoms  
30 verteilt wurde, als die Insertionssequenz aus pCB435-45 – obwohl  
auf das Ereignis der Insertion der Insertionssequenz aus  
pCB435-45 selektiert wurde (aadA Markergen befindet sich in  
pCB435-45). Die Wirksamkeit des im Rahmen dieser Erfindung be-  
schriebenen Verfahrens – nämlich der Insertion und schnellen Ver-  
35 teilung einer Insertionssequenz im Plastidengenom, ohne dass auf  
die Anwesenheit dieser selektiert würde, durch das Nutzen von  
DSBI-Enzymen und entsprechende Erkennungsstellen – konnte somit  
in diesem Beispiel für besagte Linien nachgewiesen werden.

## 111

**Beispiel 14:** Erzeugen von weiteren Masterpflanzen mit einer nicht-natürlicherweise in Plastiden vorkommenden DSB-Erkennungsregion, sowie Transformation dieser Pflanzen unter Ausnutzen des DSBI-Enzyms I-PpoI

5

14.1: Erstellen eines weiteren Vektors (pCB456-2) für die Einbringung einer nicht natürlicherweise vorkommenden Erkennungsregion für die Homing Endonuklease I-PpoI in das Plastom von Tabak

10

Ziel dieses Ansatzes war es (analog zu Beispiel 3) einen weiteren Vektor für die Plastidentransformation zu erstellen, welcher keine umfassenden Homologien zu Sequenzen im Plastidengenom besitzt.

15

Der Selektionsmarker aadA ist in diesem Plasmid unter Kontrolle eines synthetischen Promotors, der abgeleitet ist von der Konsensus Sequenz für E. coli σ<sub>70</sub> Promotoren. Als 3' Ende wurde ein Bereich stromabwärts des 3'psbA-1 Gens aus Synechocystis genutzt.

20 Im Unterschied zu dem bereits beschriebenen Vektor pCB199-3 wurde hier die DSB-Erkennungssequenz direkt stromabwärts des aadA Gens, aber stromaufwärts der 3'psbA-1 Sequenz aus Synechocystis in das Molekül eingeführt. Damit kann man nach Insertion einer Insertionssequenz mit Hilfe eines DSBI-Enzyms ein Operon erstellen. Die

25 Gene auf der Insertionssequenz können dann - optional - ohne Promotor auf der Insertionssequenz insertiert werden. Nach der Insertion gelangen entsprechende Gene der Insertionssequenz dann auch unter Kontrolle des synthetischen Promotors stromaufwärts des aadA Gens in der Masterpflanze. Dadurch kann man - optional -

30 eine Operon Struktur bestehend aus dem aadA und den nachfolgend eingeführten Genen im Plastom erzeugen.

Zur Herstellung des Plasmides pCB456-2 wurden sukzessive verschiedene Elemente in den Basisvektor pCB42-94 kloniert:

35

- Synthetischer Promotor (komplementär bp 1226-1260)
- Ribosomenbindestelle (komplementär bp 1214-1218)
- aadA Gen (komplementär bp 414-1208)
- Core-Erkennungsregion für die Homing Endonuklease I-PpoI

40

(komplementär bp 331-345)

- 3'psbA-1 aus Synechocystis (komplementär bp 19-155)

Der so erhaltene Vektor vermittelt auch in E. coli eine Spectinomycin Resistenz. Anstelle der multiplen Klonierungsstelle im Basisvektor pCB42-94 für die Plastidentransformation enthält dieser Vektor mit der Bezeichnung pCB456-2 die aufgeführten Elemente im Rahmen der Nukleinsäuresequenz mit der SEQ ID NO: 83. Auch hier

## 112

ist die gesamte MCS ersetzenende Sequenz (von SacI bis KpnI) angegeben.

- 14.2: Erstellen von überwiegend homotransplastomen Masterpflanzen, die eine nicht natürliche DSB-Erkennungssequenz enthalten

Analog zu pCB199-3 in Beispiel 4 wurde der Vektor pCB456-2 in die Plastiden von Tabak eingebracht. Abweichend von der Beschreibung 10 in Beispiel 4 wurden die erhaltenen Sprosse jedoch auf Anzuchtmedium, welches 30 g/L Sacharose enthielt (anstelle der in Beispiel 4 angegebenen 10 g/L), kultiviert. Die resultierenden Pflanzen wurden mit CB456NTH bezeichnet. Mittels Southern-Hybridisierung wurden von den Spectinomycin resistenten Pflanzen, die nach der 15 Transformation erhalten wurden, 2 Linien (CB456NTH-1 und -15, vgl. Fig. 12) identifiziert, die die Insertionssequenz aus pCB456-2 in ihr Plastom eingebaut haben. In dem Southern-Experiment wurde eine Sonde genutzt, welche gegen ein Fragment der 16SrDNA gerichtet war (vgl. oben Beispiel 13.2). Diese Sonde war geeignet, um aus mit EcoRI verdauter DNA, die dem Wildtyp entspricht, ein Fragment von etwa 3,1 kb nachzuweisen. Im Gegensatz dazu wurde ein etwa 1,7 kb großes Fragment detektiert, wenn in 20 die entsprechenden Plastomkopien die Insertionssequenz aus pCB456-2 eingebaut worden war.

- 25 14.3: Erstellen eines Transformationsvektors für artifizielles Homing in den Masterpflanzen CB456NTH

Zunächst wurde eine Operonstruktur bestehend aus den Elementen 30 RBS - nptII (codierend für ein Enzym, welches Kanamycin Resistenz vermittelt) - RBS - I-PpoI (codierend für ein DSBI-Enzym) erstellt. Diese Kassette wurde mit BstXI Schnittstellen umgeben, die nach Einwirkung des Enzyms BstXI DNA-Enden erzeugen, die kompatibel zu den durch das Enzym I-PpoI erzeugten DNA-Enden sind. 35 Der resultierende Vektor (Rückgrat entspricht dem des pBluescript) wurde mit pCB528-2 bezeichnet, enthält ein Insert gemäß SEQ ID NO: 84 mit folgenden Elementen:

- RBS (Pos. 28-32)
- 40 - nptII (Pos. 27-840)
- RBS (Pos. 849-853)
- Gen codierend für I-PpoI (Pos. 859-1350)

Anschließend wurde mit BstXI das 1360 bp große Fragment aus 45 pCB528-2 heraus geschnitten und in die I-PpoI Schnittstelle im Vektor pCB456-2 ligiert. Aus den erhaltenen Klonen nach Transformation der Ligationsprodukten in E. coli wurden solche ausge-

113

wählt, die eine Kanamycin Resistenz besaßen. Dadurch wurde sichergestellt, dass besagtes Insert in dem entsprechenden Klon derart in den Vektor insertiert war, dass die nptII und I-PpoI Kassette in der selben Orientierung vorlagen, wie die aadA-Kasse-  
5 sette. Dies wurde auch durch das dem Fachmann bekannte Verfahren der Restriktionsanalyse verifiziert. Der entsprechende Vektor erhielt die Bezeichnung pCB535-11.

14.4: Transformation von pCB535-11 in Masterpflanzen CB456NTH  
10 Analog zu pCB456-2 in Beispiel 14.2 wurde pCB535-11 auf Goldpartikel aufgebracht und anschließend mittels Partikelkanone in Plastiden der Masterpflanze CB456NTH-1 eingebracht.

Ein Teil der Explantate wurde ohne jeglichen Selektionsdruck auf  
15 Regenerationsmedium inkubiert. Entstandene Pflanzen wurden auf Anzuchtmedium (wieder ohne Selektionsdruck überführt). Anschließend werden die Pflanzen mittels PCR in Hinblick auf die Anwesenheit der RBS - nptII - RBS - I-PpoI Kassette analysiert.

20 Ein anderer Teil der Explantate wurde auf Regenerationsmedium zugleich 15 bzw. 30 mg/L Kanamycin inkubiert. Nach je 2 Wochen werden die Pflanzen auf frisches Regenerationsmedium umgesetzt und die Kanamycinkonzentration sukzessive bis 50 bzw. 80 mg/L erhöht.

25

30

35

40

45

## Patentansprüche

1. Verfahren zur Integration einer DNA-Sequenz in die plastidäre  
5 DNA einer mehrzelligen Pflanze oder von dieser abgeleiteten  
Zelle und zur Selektion überwiegend homotransplastomer Zellen  
oder Pflanzen, dadurch gekennzeichnet, dass
    - a) die plastidären DNA-Moleküle besagter mehrzelliger  
10 Pflanze mindestens eine Erkennungssequenz zur gezielten  
Induktion von DNA-Doppelstrangbrüchen enthalten und
    - b) mindestens ein Enzym geeignet zur Induktion von  
15 DNA-Doppelstrangbrüchen an der Erkennungssequenz zur  
gezielten Induktion von DNA-Doppelstrangbrüchen sowie  
mindestens ein Transformationskonstrukt umfassend eine  
Insertionssequenz in mindestens einem Plastid besagter  
mehrzelligen Pflanze oder von dieser abgeleiteten Zelle  
zusammengebracht werden, und
  - 20 c) die Induktion von DNA-Doppelstrangbrüchen an den  
Erkennungssequenzen zur gezielten Induktion von DNA-  
Doppelstrangbrüchen erfolgt, und
  - 25 d) die Insertionssequenz in die plastidäre DNA insertiert,  
wobei die Funktionalität der Erkennungssequenz zur ge-  
zielten Induktion von DNA-Doppelstrangbrüchen deaktiviert  
wird, so dass besagte Erkennungssequenz nicht mehr durch  
das Enzym geeignet zur Induktion von DNA-Doppelstrang-  
brüchen geschnitten werden kann, und
  - 30 e) Pflanzen oder Zellen isoliert werden, bei denen die  
Insertionssequenz in die plastidären DNA-Moleküle  
insertiert wurde.
- 35
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die  
Erkennungssequenz zur gezielten Induktion von DNA-Doppel-  
strangbrüchen sowie die Insertionssequenz zumindest ein-  
seitig von Sequenzen flankiert sind, die eine ausreichende  
40 Länge und ausreichende Homologie aufweisen, um eine homologe  
Rekombination untereinander zu gewährleisten.
  - 45 3. Verfahren nach einem der Ansprüche 1 oder 2, dadurch gekenn-  
zeichnet, dass das Transformationskonstrukt mindestens eines  
der Elemente beinhaltet ausgewählt aus der Gruppe bestehend  
aus

## 115

- i) Expressionskassette für ein Enzym, geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen
  - 5 ii) Positiven Selektionsmarkern
  - iii) Negativen Selektionsmarkern
  - 10 iv) Reportergenen
  - v) Replikationsursprünge
  - vi) Multiple Klonierungsregionen
  - 15 vii) "Border"-Sequenzen für Agrobakterium-Transfektion
  - viii) Sequenzen, die eine homologe Rekombination bzw. Insertion in das Genom eines Wirtsorganismus ermöglichen
  - 20
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass ein Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen ausgewählt ist aus der Gruppe bestehend aus Restriktionsendonukleasen, Homing-Endonukleasen, Chimäre Nukleasen, mutierte Restriktions- oder Homing-Endonukleasen und RNA-Proteinpartikel abgeleitet von mobilen Intronen der Gruppe II, sowie Fusionsproteinen der vorgenannten Proteine mit Plastidenlokalisationssequenzen.
- 25
5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass das Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen ausgewählt ist aus der Gruppe der Homing-Endonukleasen bestehend aus F-SceI, F-SceII, F-SuVI, F-TevI, F-TevII, I-AmAI, I-AnII, I-CeuI, I-CeuAIIIP, I-ChuI, I-CmoI, I-CpaI, I-CpaII, I-CreI, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIVP, I-CsmI, I-CvuI, I-CvuAIP, I-DdII, I-DirI, I-DmoI, I-HspNIP, I-LlaI, I-MsoI, I-NaaI, I-NanI, I-NclIP, I-NgrIP, I-NitI, I-NjaI, I-Nsp236IP, I-PakI, I-PboIP, I-PcuIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PpbIP, I-PpoI, I-SPBetaIP, I-ScaI, I-SceI, I-SceII, I-SceIII, I-SceIV, I-SceV, I-SceVI, I-SceVII, I-SexIP, I-SneIP, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquIP, I-Ssp6803I, I-StPhiJP, I-StPhiST3P, I-StPhiS3bP, I-TdeIP, I-TevI, I-TevII,
- 30
- 35
- 40
- 45

## 116

I-TevIII, I-UarAP, I-UarHGPA1P, I-UarHGPA13P, I-VinIP,  
I-ZbiIP, PI-MtuI, PI-MtuHIP, PI-MtuHIIIP, PI-PfuI, PI-PfuII,  
PI-PkoI, PI-PkoII, PI-PspI, PI-Rma43812IP, PI-SPBetaIP,  
PI-SceI, PI-TfuI, PI-TfuII, PI-ThyI, PI-TliI, PI-TliII

5 sowie Fusionsproteinen der vorgenannten mit plastidären Lokalisationssequenzen.

6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass das Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen ausgewählt ist aus der Gruppe der Homing-Endonukleasen bestehend aus den Enzymen gemäß SEQ ID NO: 5, 12, 14 und 71.
- 10 15 7. Verfahren nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, dass das Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen unter Verwendung einer Expressionskassette exprimiert wird, die eine für das besagte Enzym kodierende Nukleinsäuresequenz beinhaltet.
- 15 20 8. Expressionkassette enthaltend ein Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen gemäß Definition einer der Ansprüche 4 bis 6 unter Kontrolle eines in pflanzlichen Plastiden funtionellen Promotors.
- 25 30 9. Expressionkassette enthaltend ein Fusionsprotein aus einer Plastidentranslokalisationssequenz und einem Enzym geeignet zur Induktion von DNA-Doppelstrangbrüchen an der Erkennungssequenz zur gezielten Induktion von DNA-Doppelstrangbrüchen gemäß Definition einer der Ansprüche 4 bis 6 unter Kontrolle eines im pflanzlichen Zellkern funktionellen Promotors.
- 35 40 10. Mehrzellige Pflanze erhalten nach einem Verfahren gemäß einem der Ansprüche 1 bis 7.
11. Mehrzellige Pflanze enthaltend eine Expressionskassette gemäß Anspruch 8 insertiert in die plastidäre DNA.
12. Mehrzellige Pflanze enthaltend eine Expressionskassette gemäß Anspruch 9 insertiert in die nukleäre DNA.
13. Mehrzellige Pflanze nach einem der Ansprüche 10 oder 12 ausgewählt aus der Gruppe bestehend aus *Arabidopsis thaliana*, Tabak, *Tagetes*, Weizen, Roggen, Gerste, Hafer, Raps, Mais,

117

- Kartoffel, Zuckerrübe, Soja, Sonnenblume, Kürbis oder Erdnuss.
14. Zellkulturen, Organe, Gewebe, Teile oder transgenes Vermehrungsgut abgeleitet von einer mehrzellige Pflanze nach den Ansprüchen 10 bis 13.
15. Verwendung eines Organismus nach einem der Ansprüche 10 bis 13 oder von diesem abgeleitete Zellkulturen, Organe, Gewebe, Teile oder transgenes Vermehrungsgut nach Anspruch 14 als Nahrungs-, Futtermittel oder Saatgut oder zur Herstellung von Pharmazeutika oder Feinchemikalien.
16. DNA-Konstrukt umfassend mindestens eine Nukleinsäuresequenz sowie Intronsequenzelemente, die befähigt sind, in einer von besagtem DNA-Konstrukt abgeleiteten Ribonukleinsäuresequenz die Deletion des für besagte Nukleinsäuresequenz kodierenden Ribonukleinsäurefragmentes zu gewährleisten, wobei besagte Nukleinsäuresequenz in Bezug auf besagte Intronsequenzelemente heterolog ist.
17. DNA-Konstrukt nach Anspruch 16, wobei die Nukleinsäuresequenz zumindesten von einer Spleißakzeptorsequenz und einer Spleißdonorsequenz flankiert ist.
18. DNA-Konstrukt nach Anspruch 16 oder 17, wobei das DNA-Konstrukt am 5'- und 3'-Ende Sequenzen H1 bzw. H2 umfaßt, die eine ausreichende Länge und Homologie zu plastidären Sequenzen H1' bzw. H2' aufweisen, um eine homologe Rekombination zwischen H1 und H1' bzw. H2 und H2' und damit eine Insertion der von H1 und H2 flankierten Sequenz in das Plastom zu gewährleisten.
19. Transgene Plastiden-DNA umfassend mindestens eine Nukleinsäuresequenz sowie Intronsequenzelemente, die befähigt sind, in einer von besagter transgenen Plastiden-DNA abgeleiteten Ribonukleinsäuresequenz die Deletion des für besagte Nukleinsäuresequenz kodierenden Ribonukleinsäurefragmentes zu gewährleisten, wobei die besagte Nukleinsäuresequenz in Bezug auf besagte Intronsequenzelemente heterolog ist.
20. Transgene Plastiden-DNA nach Anspruch 19, wobei die Nukleinsäuresequenz zumindesten von einer Spleißakzeptorsequenz und einer Spleißdonorsequenz flankiert ist.
21. Transgene Pflanze, umfassend eine transgene Plastiden DNA gemäß Anspruch 20.

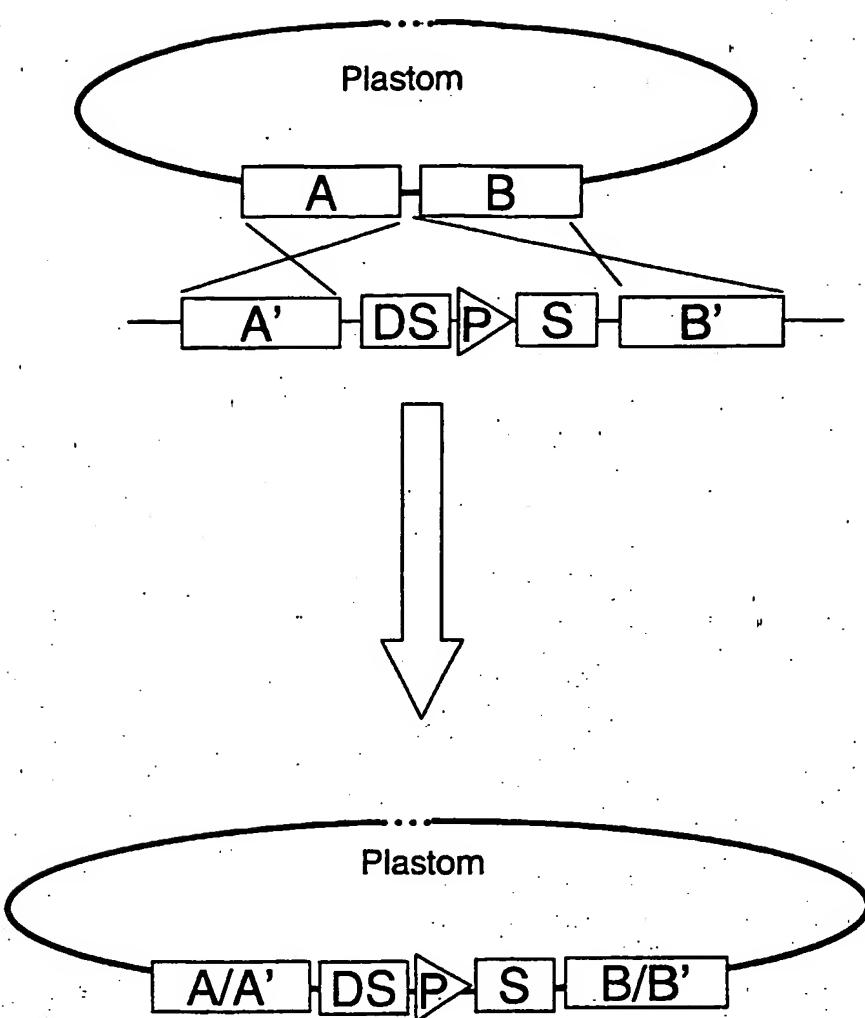


Fig. 1

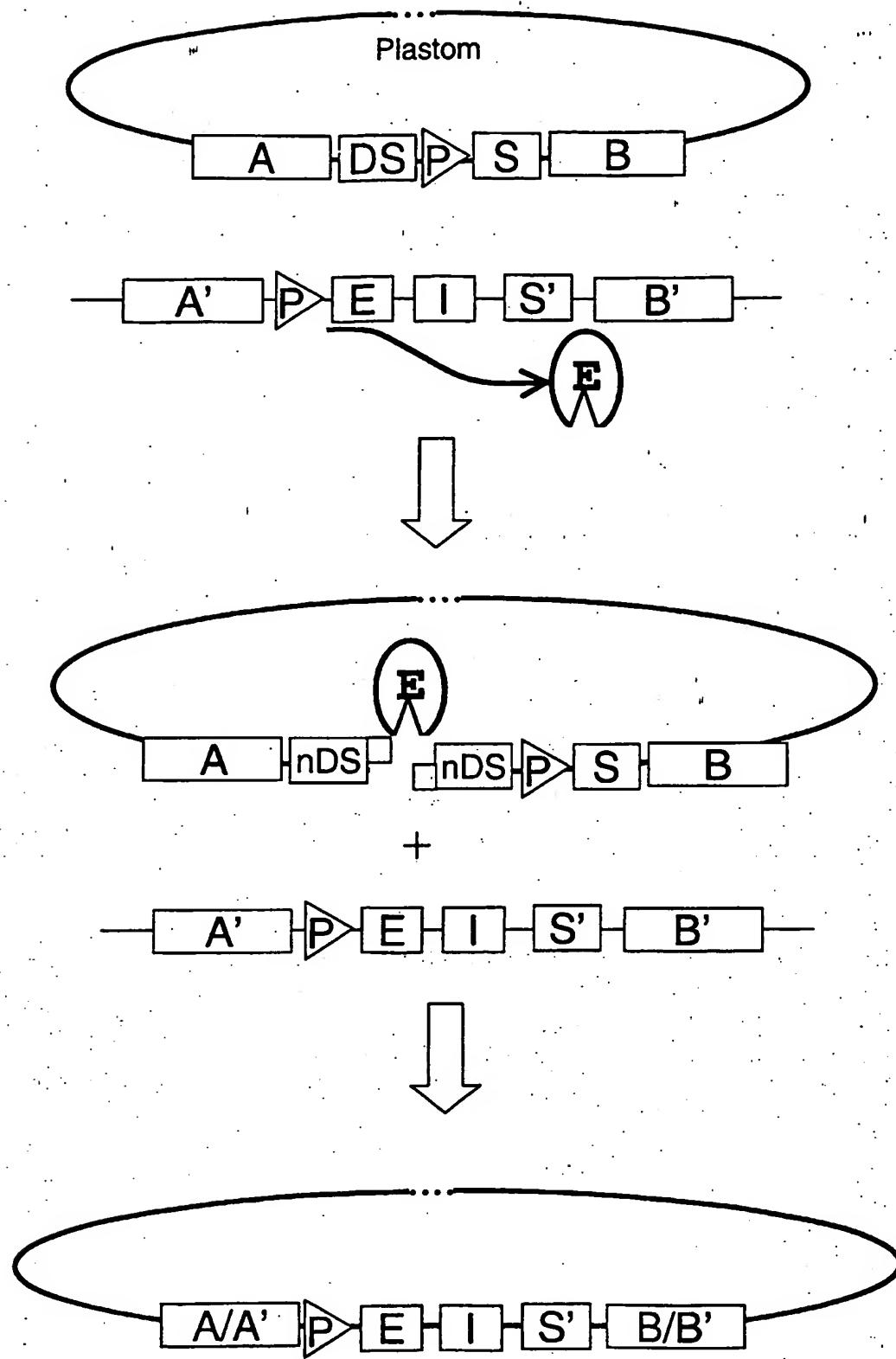


Fig. 2A

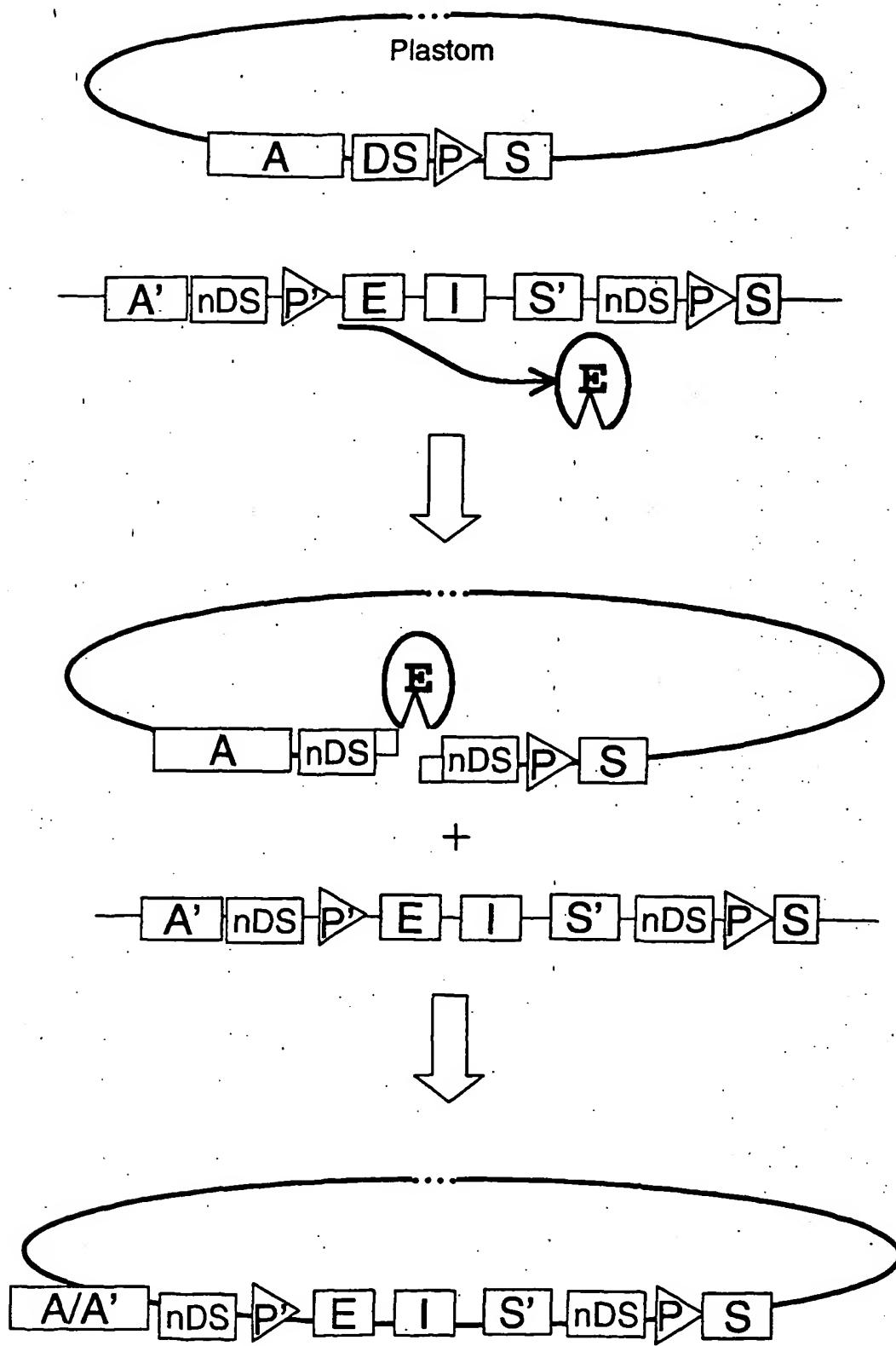


Fig. 2B

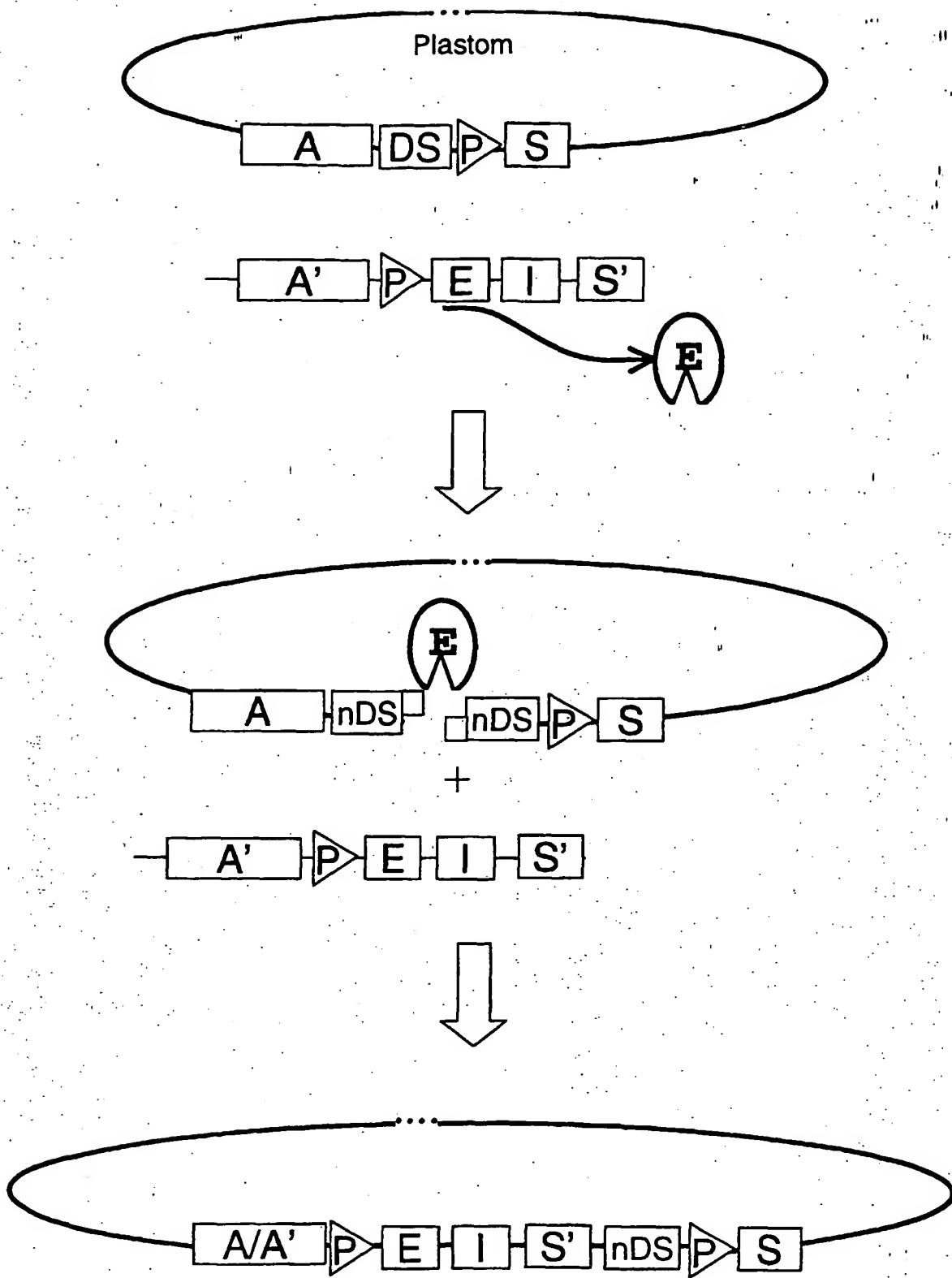


Fig. 2C

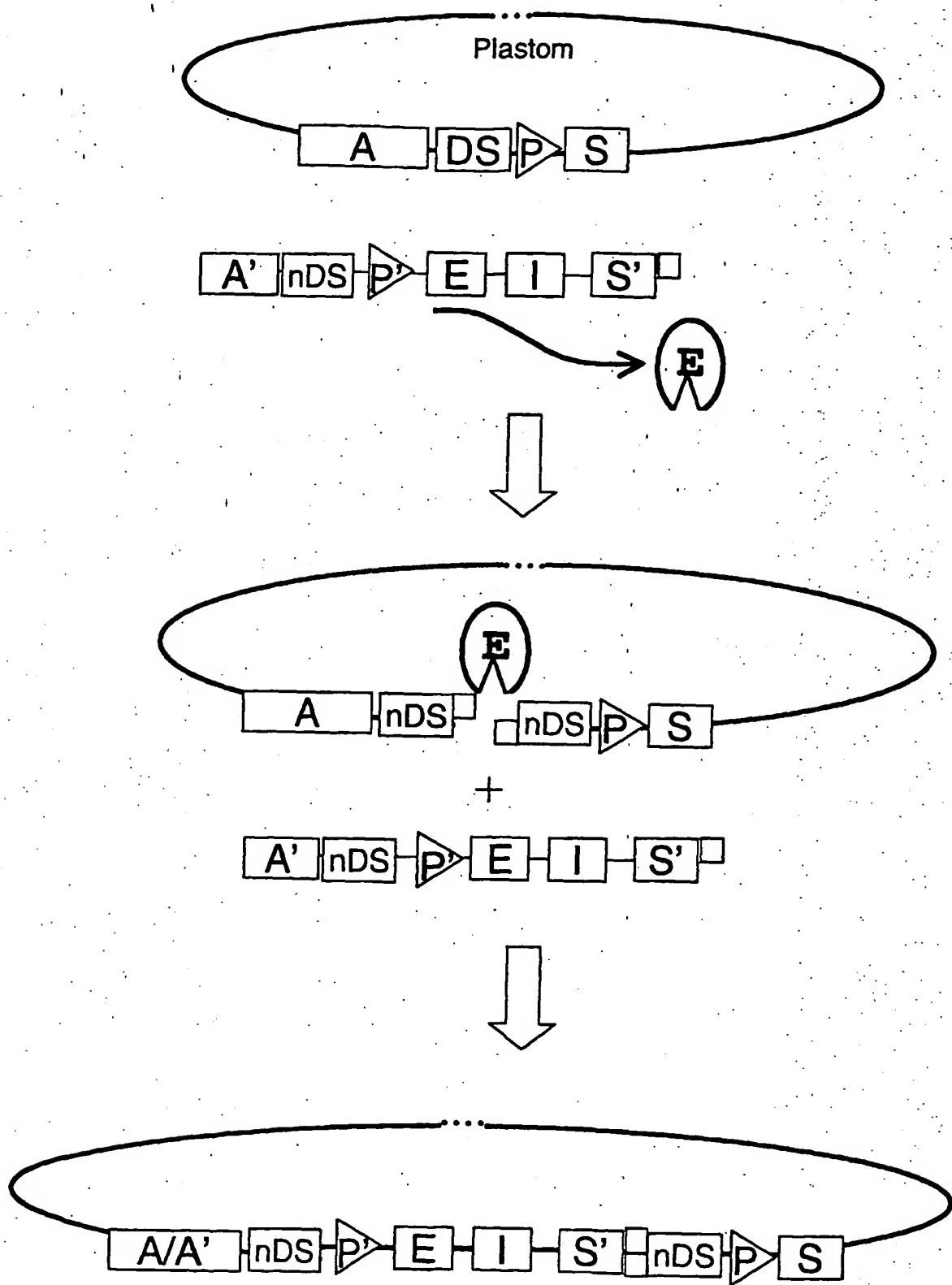


Fig. 2D

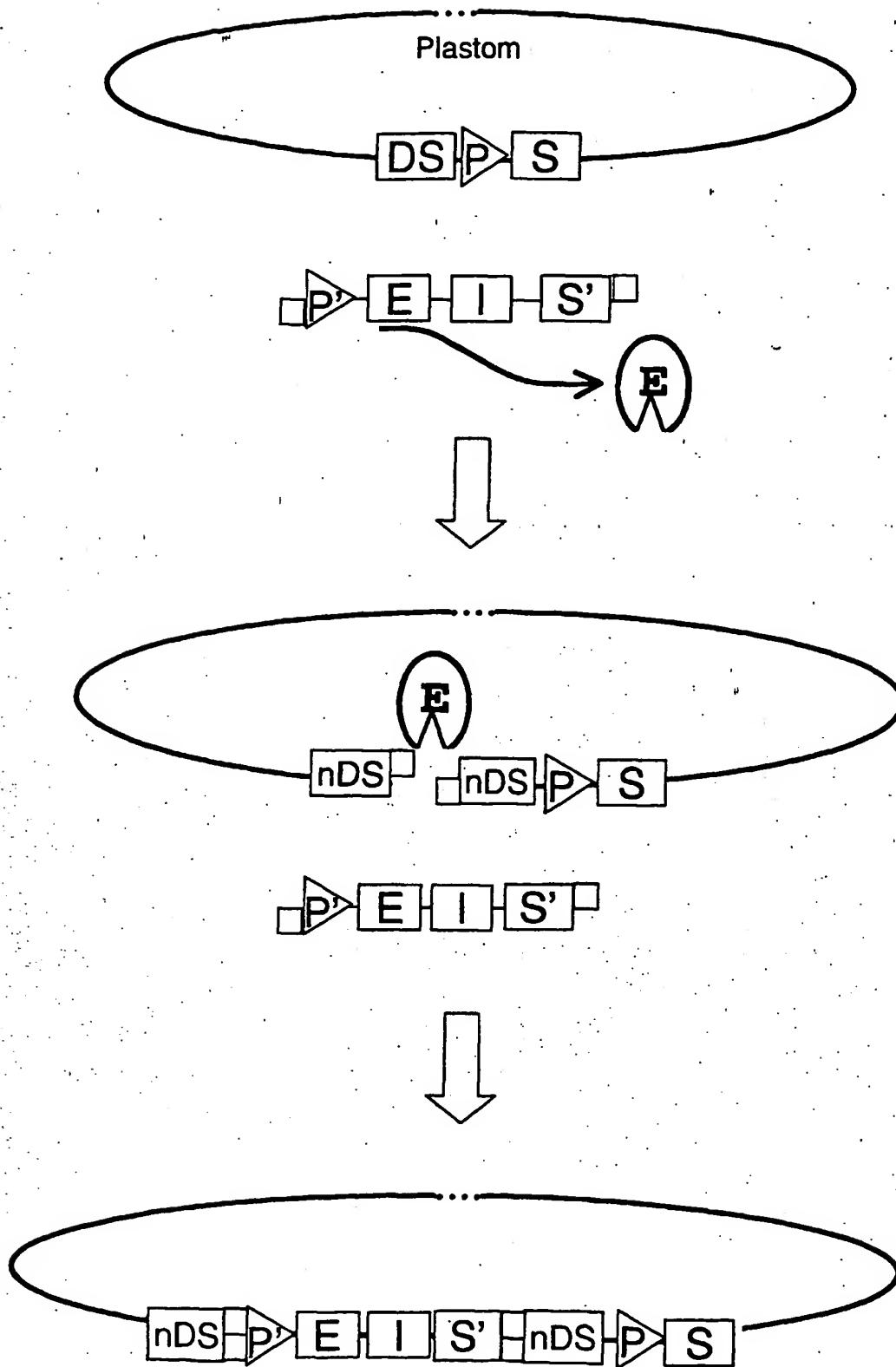


Fig. 2E

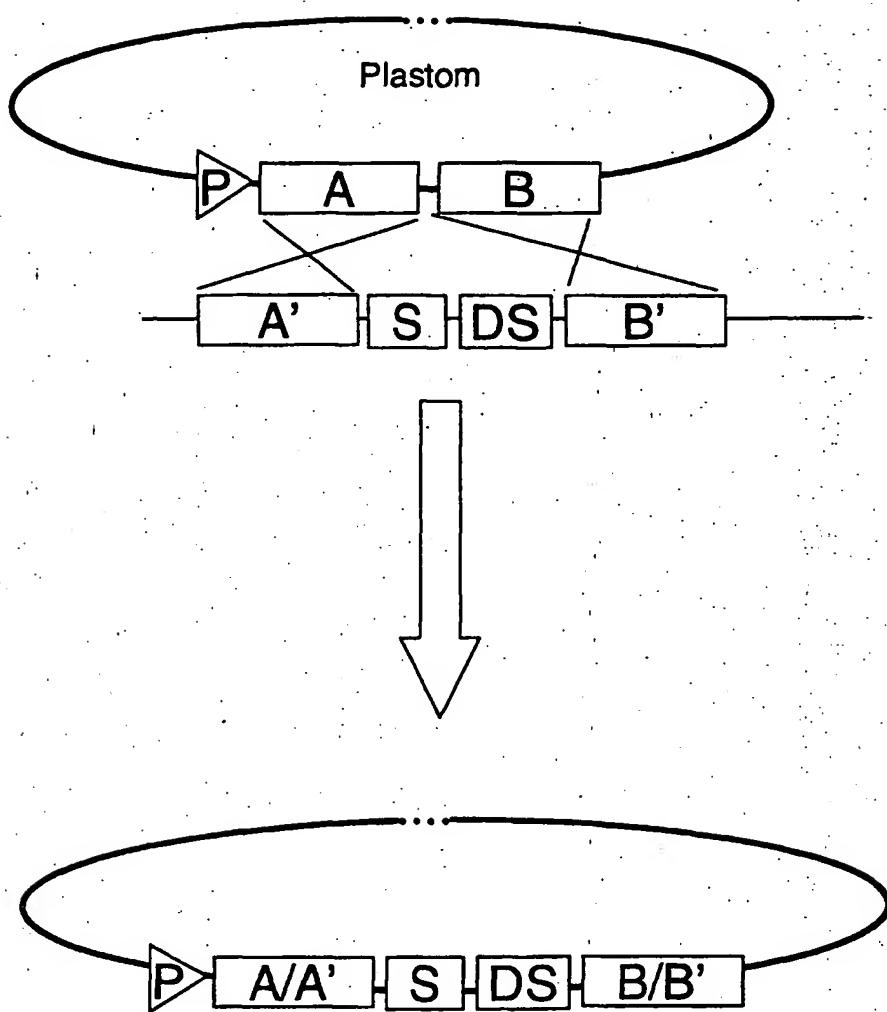


Fig. 3

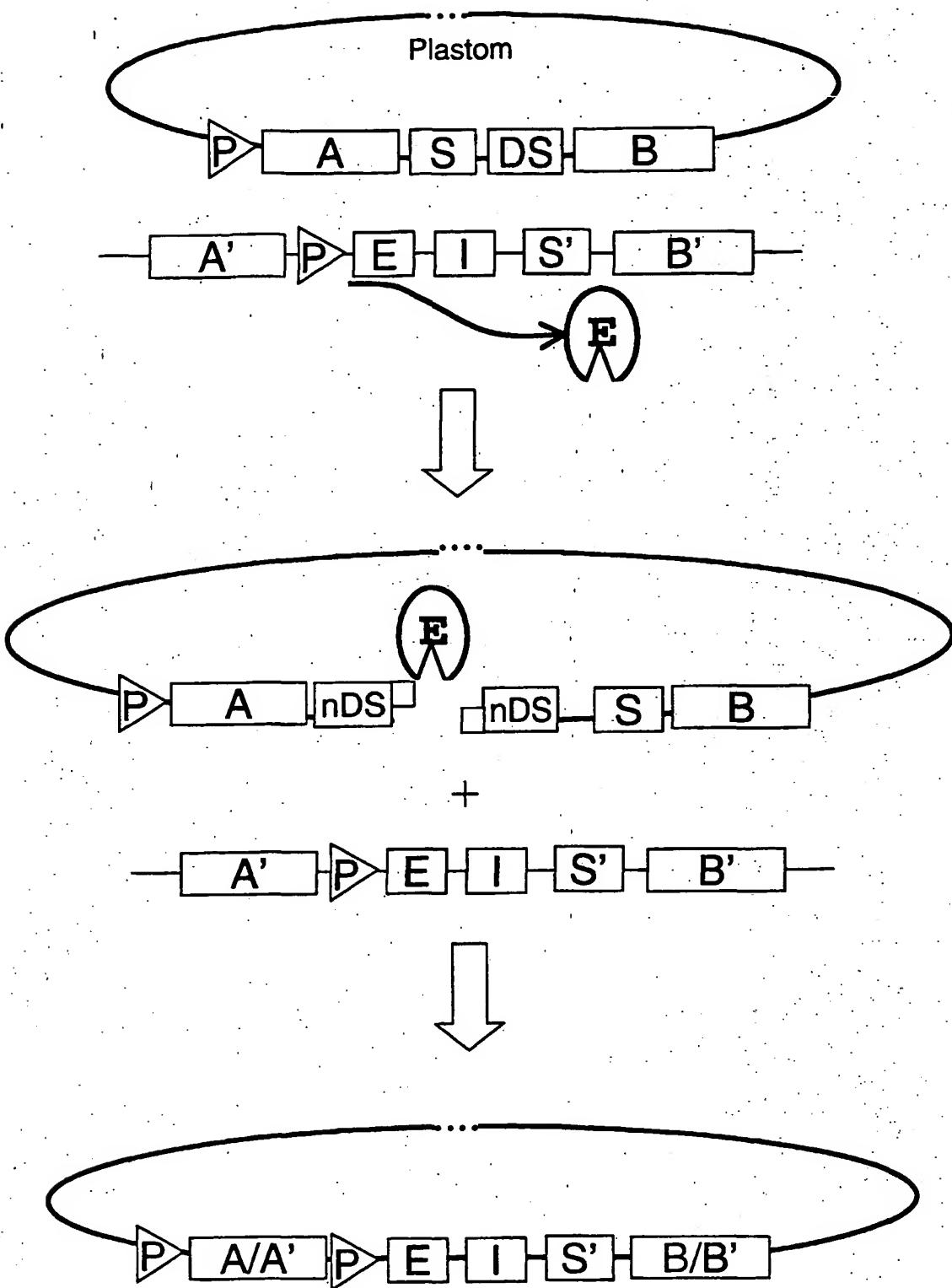


Fig. 4A

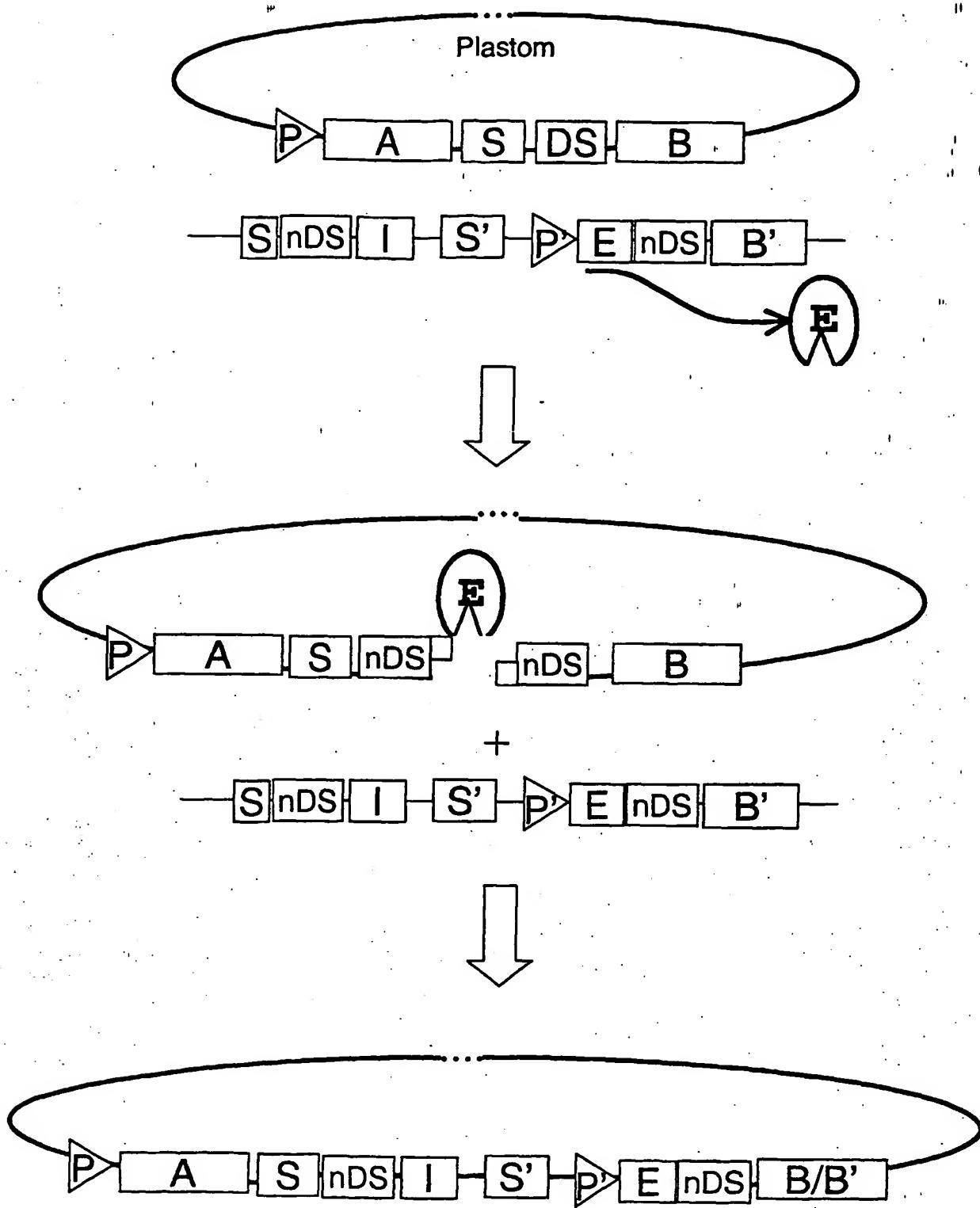


Fig. 4B

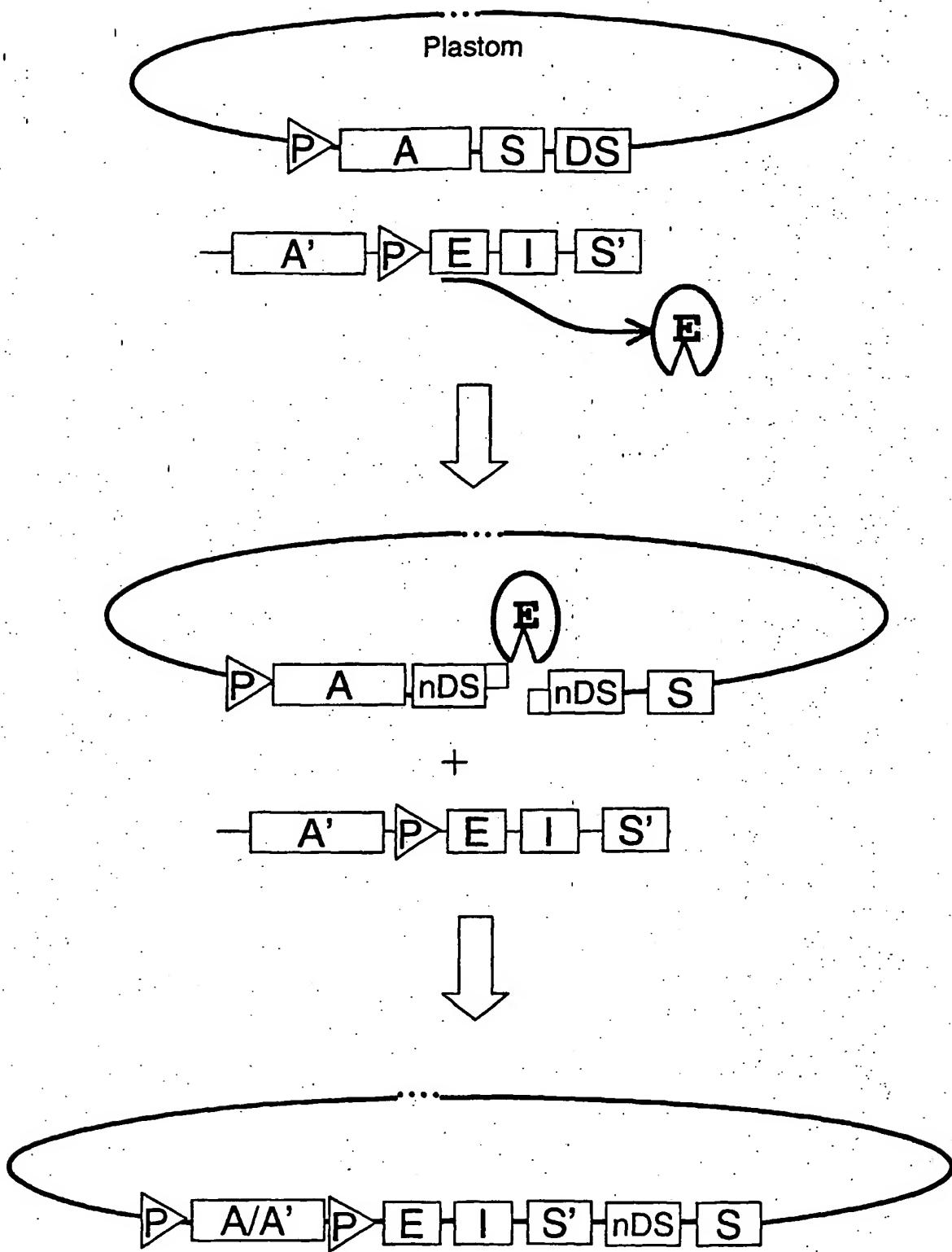


Fig. 4C

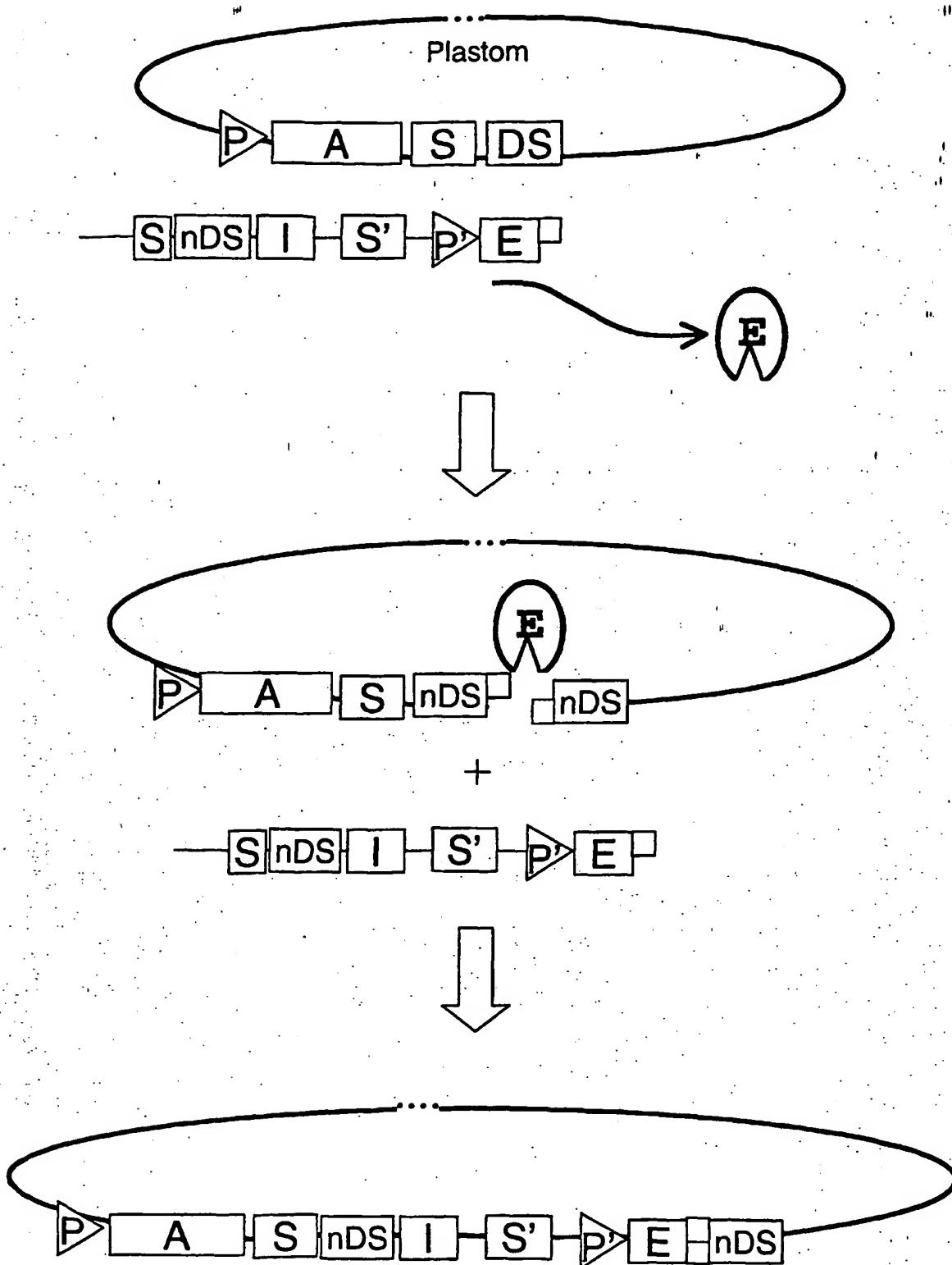


Fig. 4D

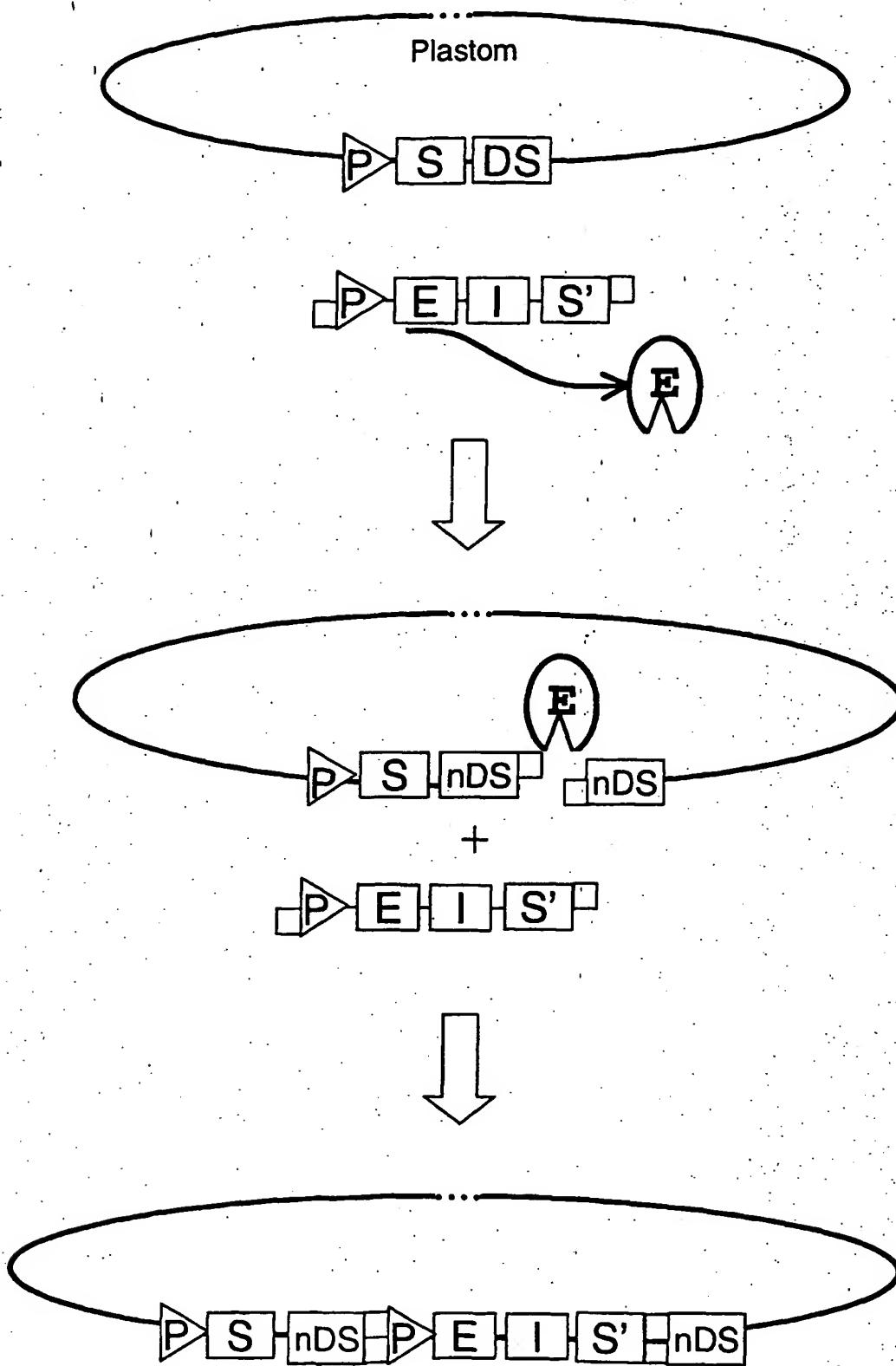


Fig. 4E

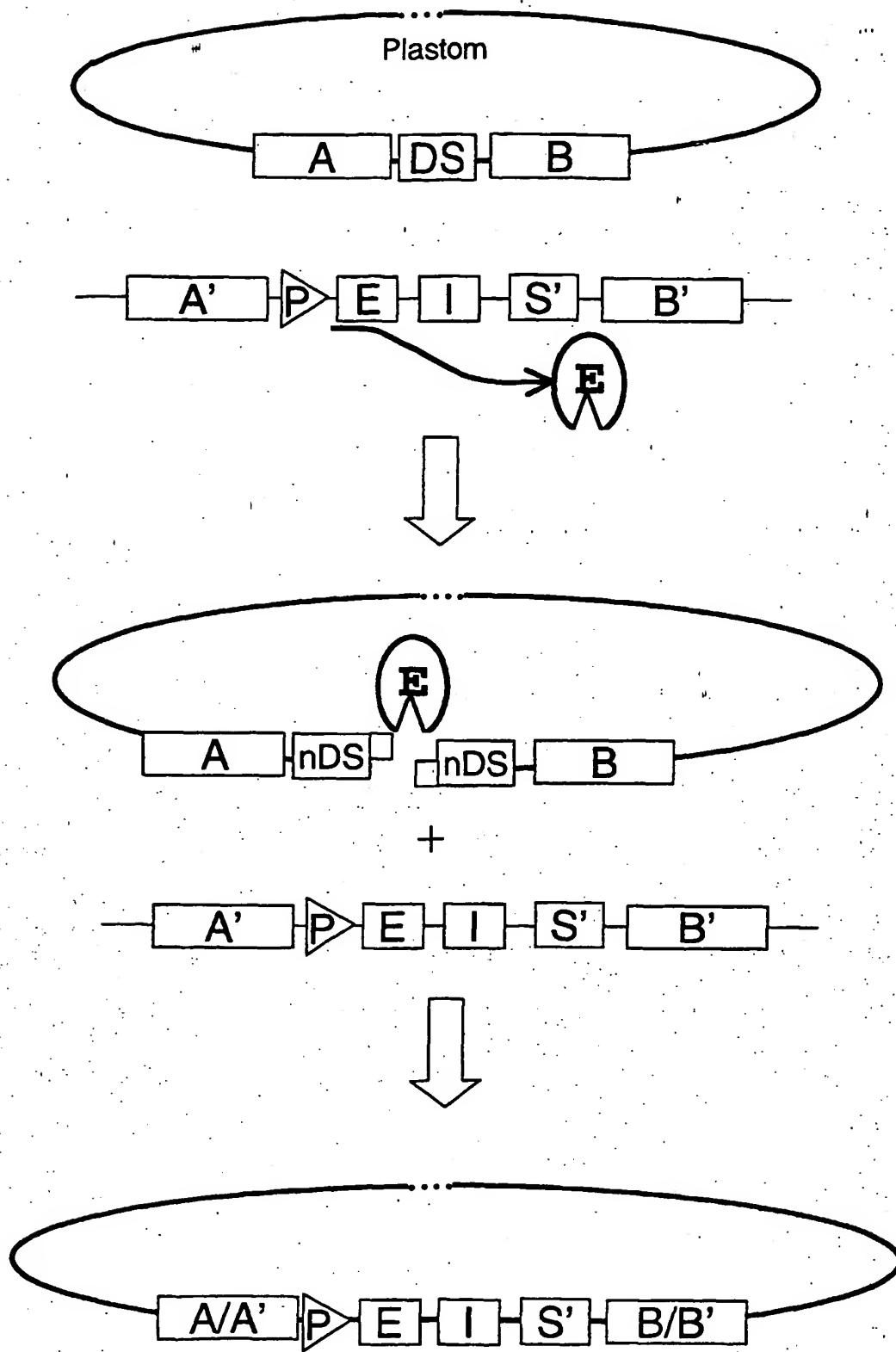


Fig. 5A

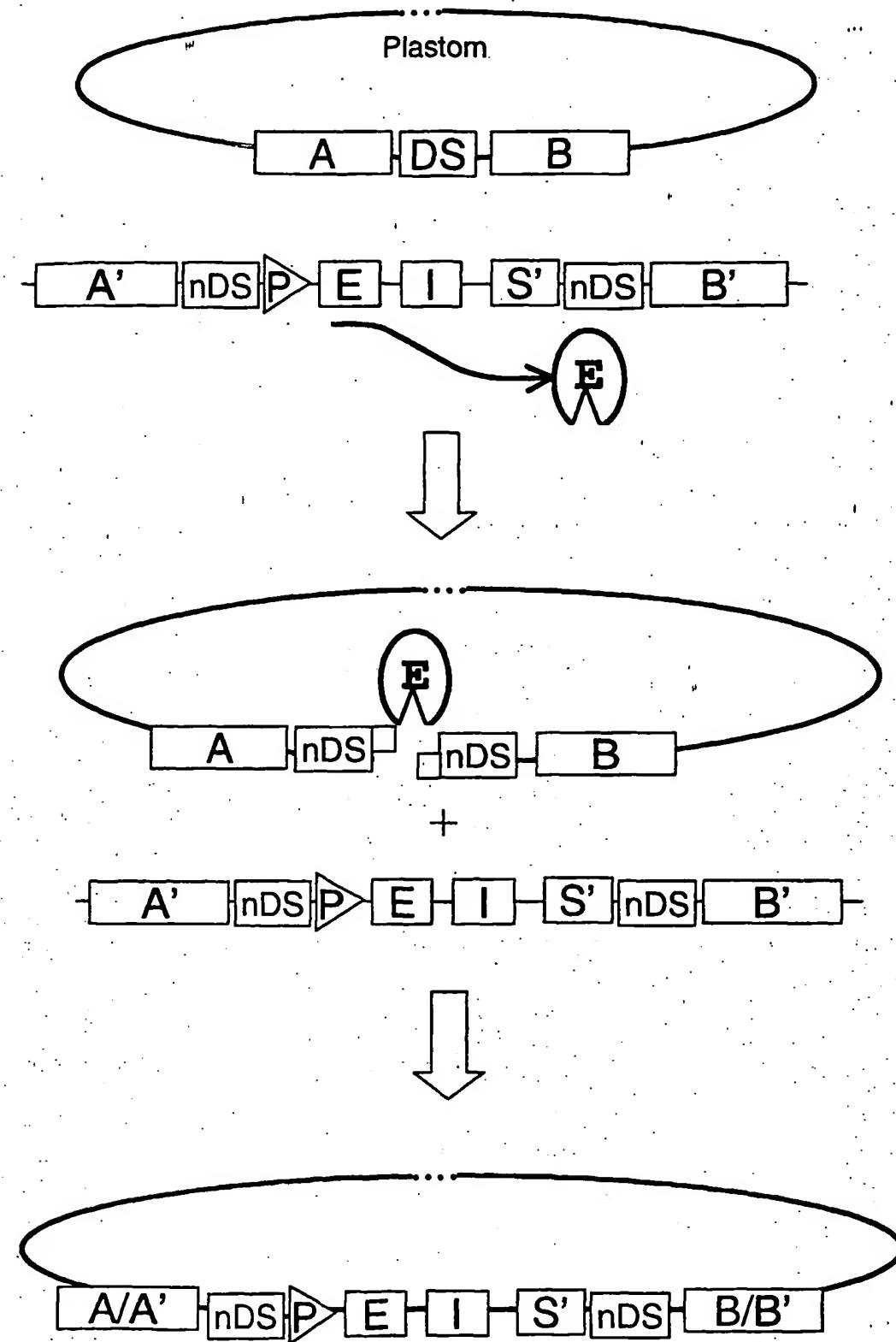


Fig. 5B

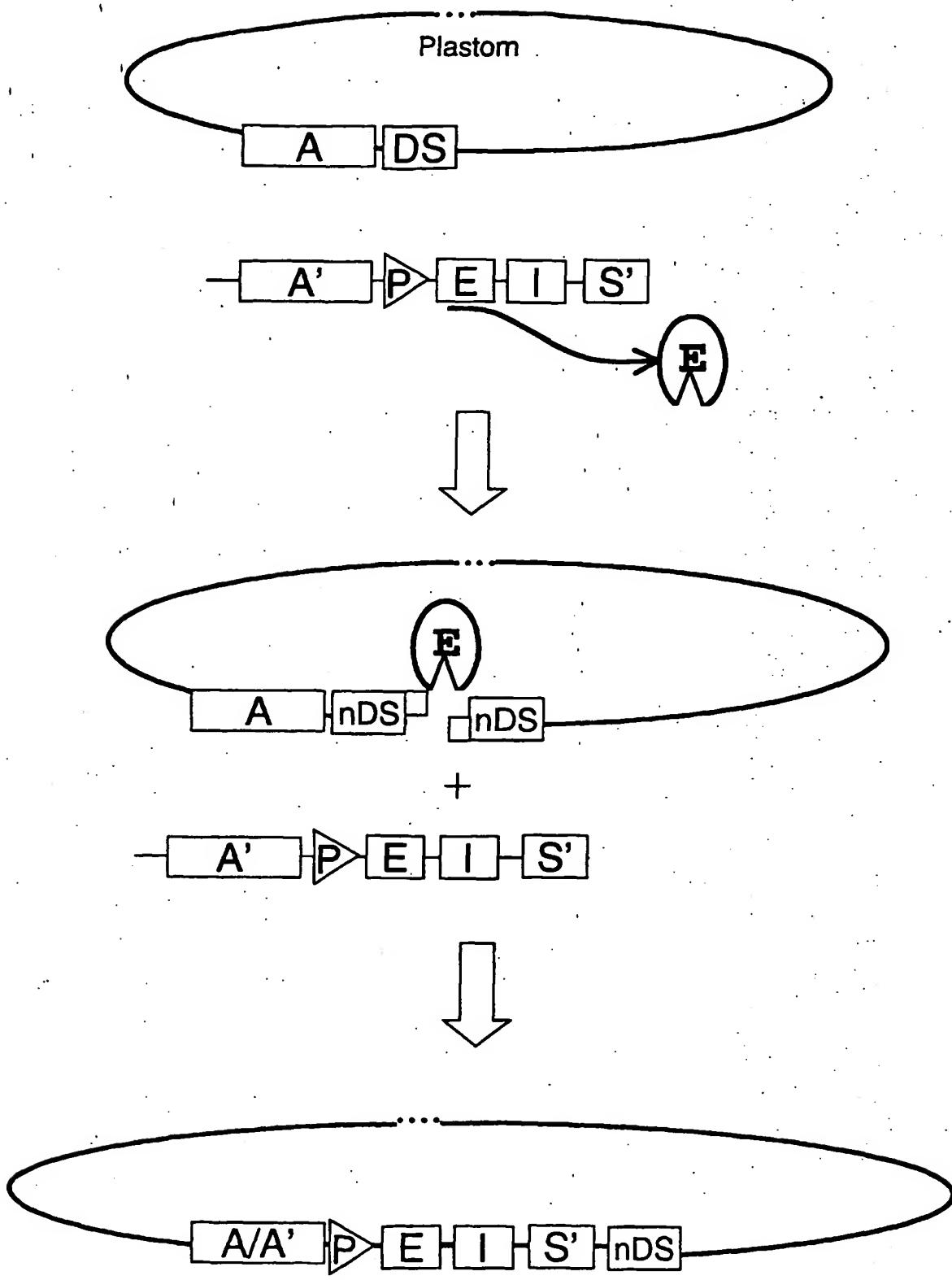


Fig. 5C

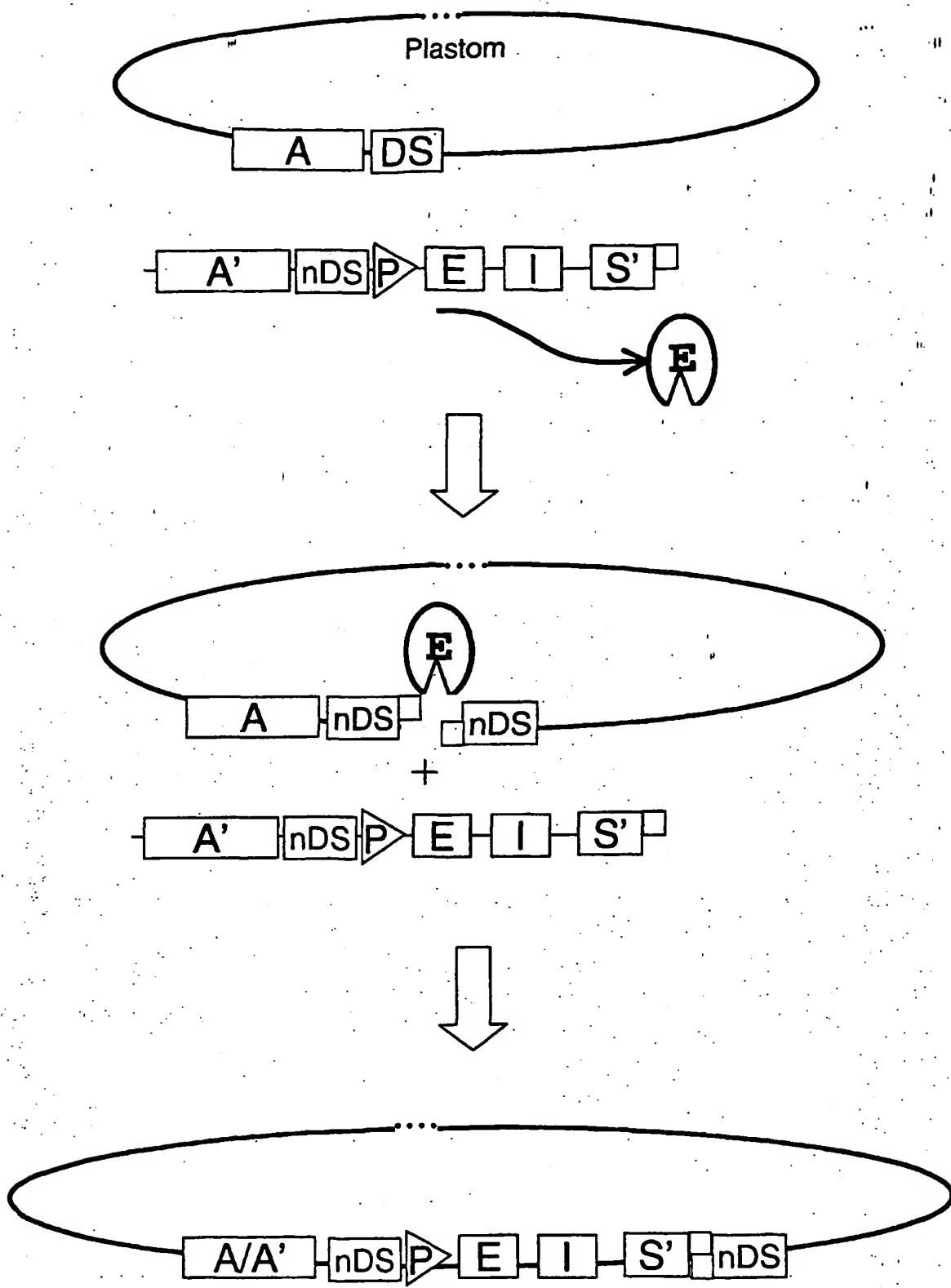


Fig. 5D

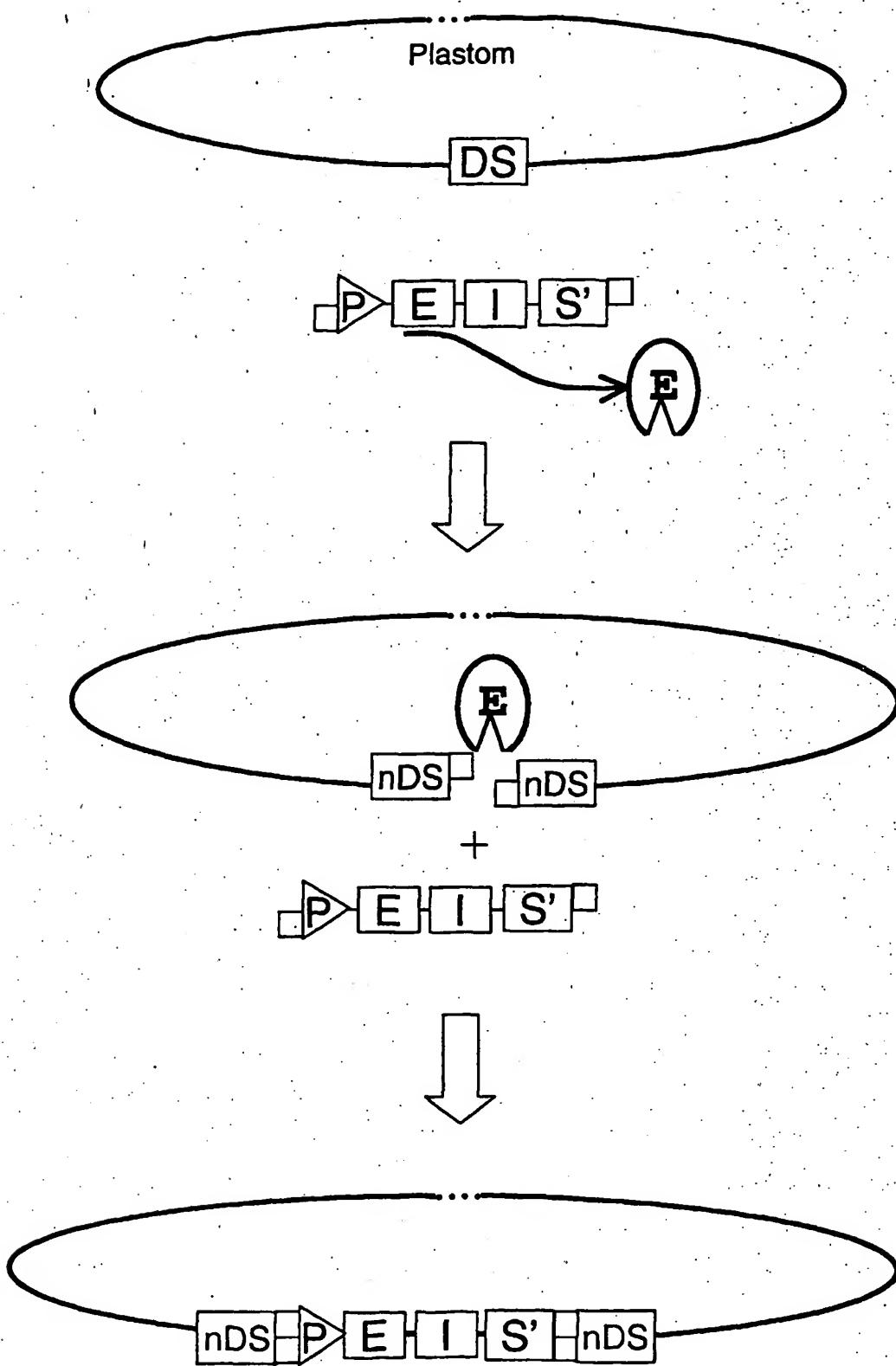


Fig. 5E

18/32

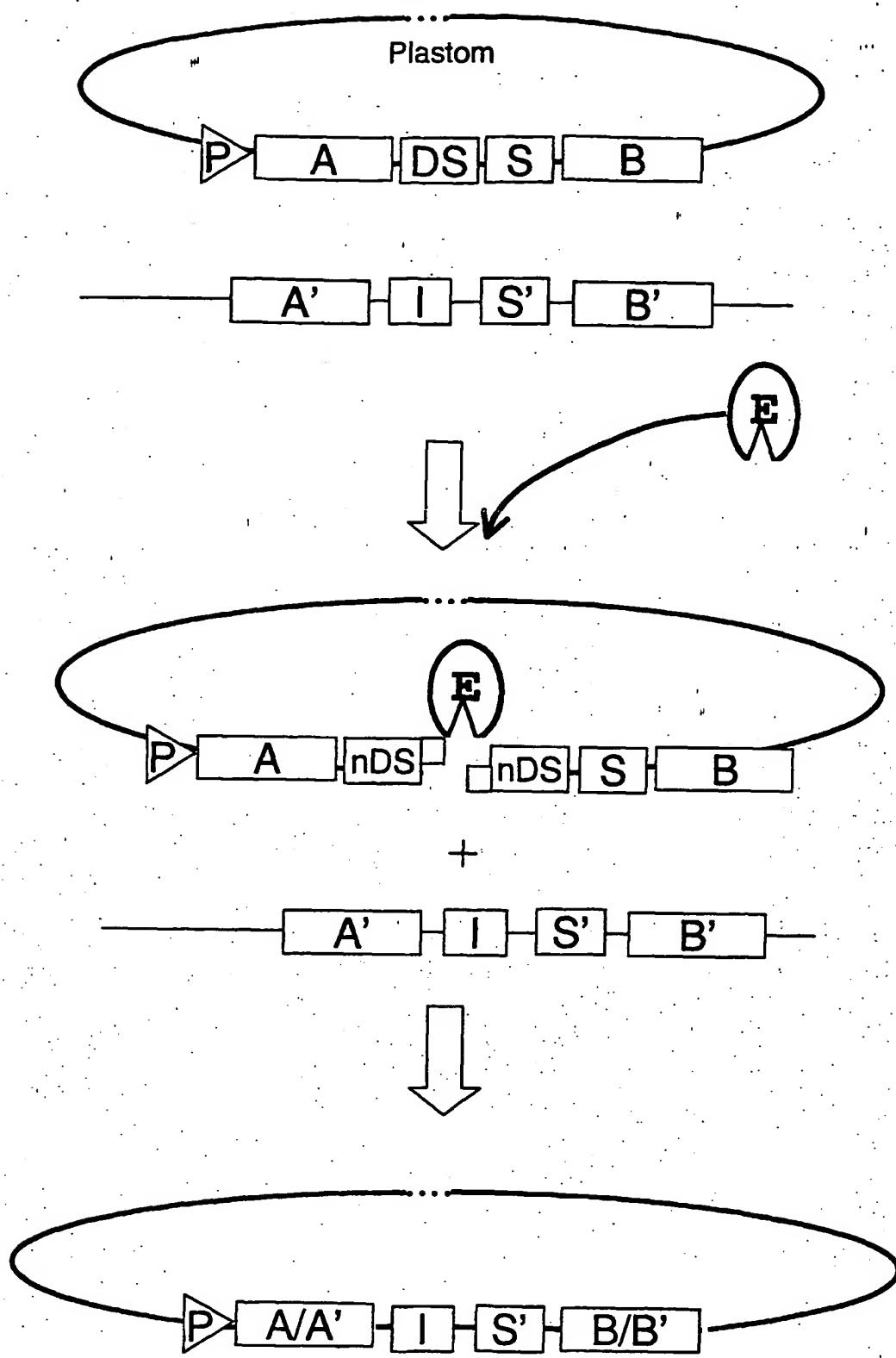


Fig. 6A

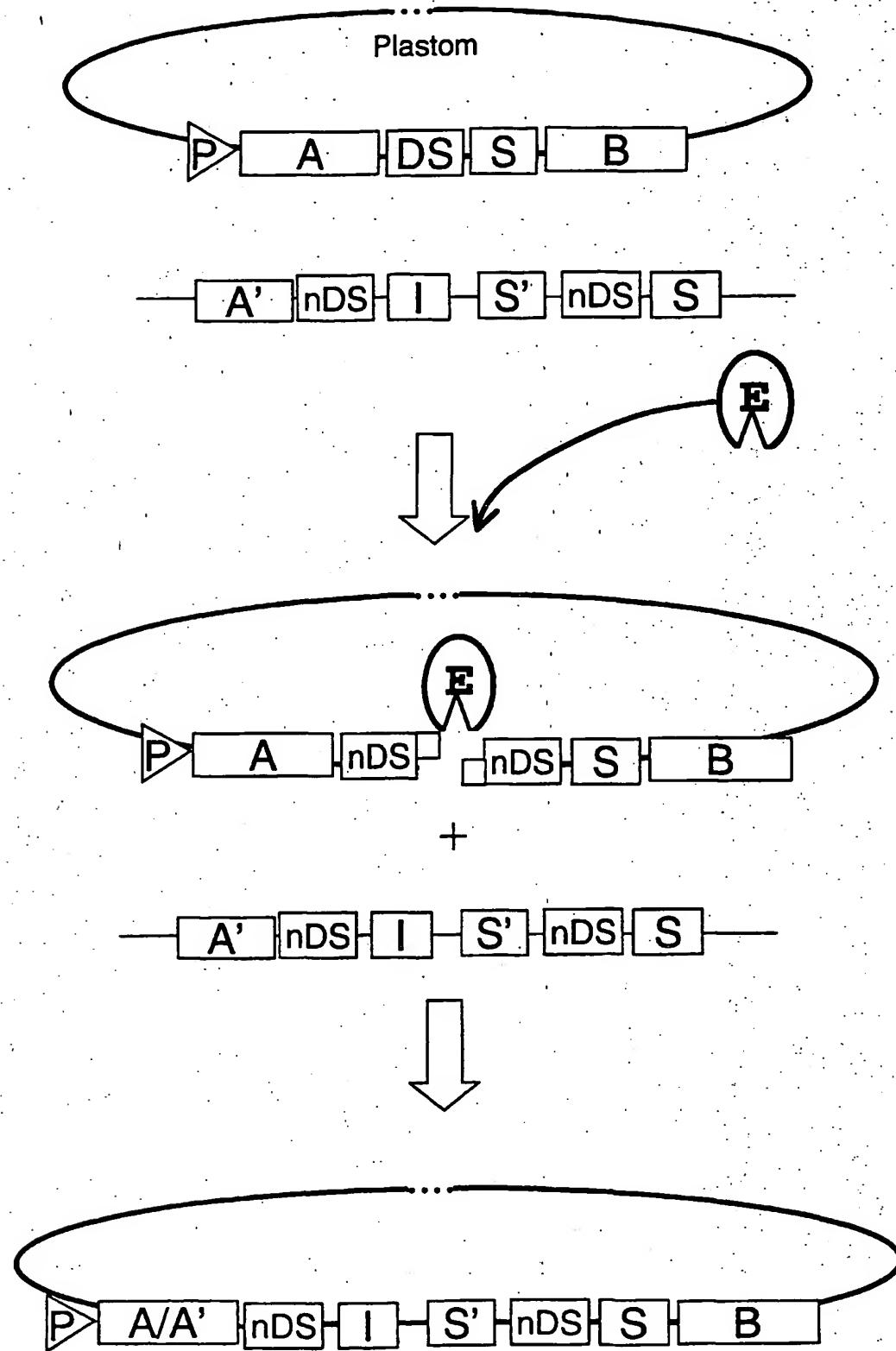


Fig. 6B

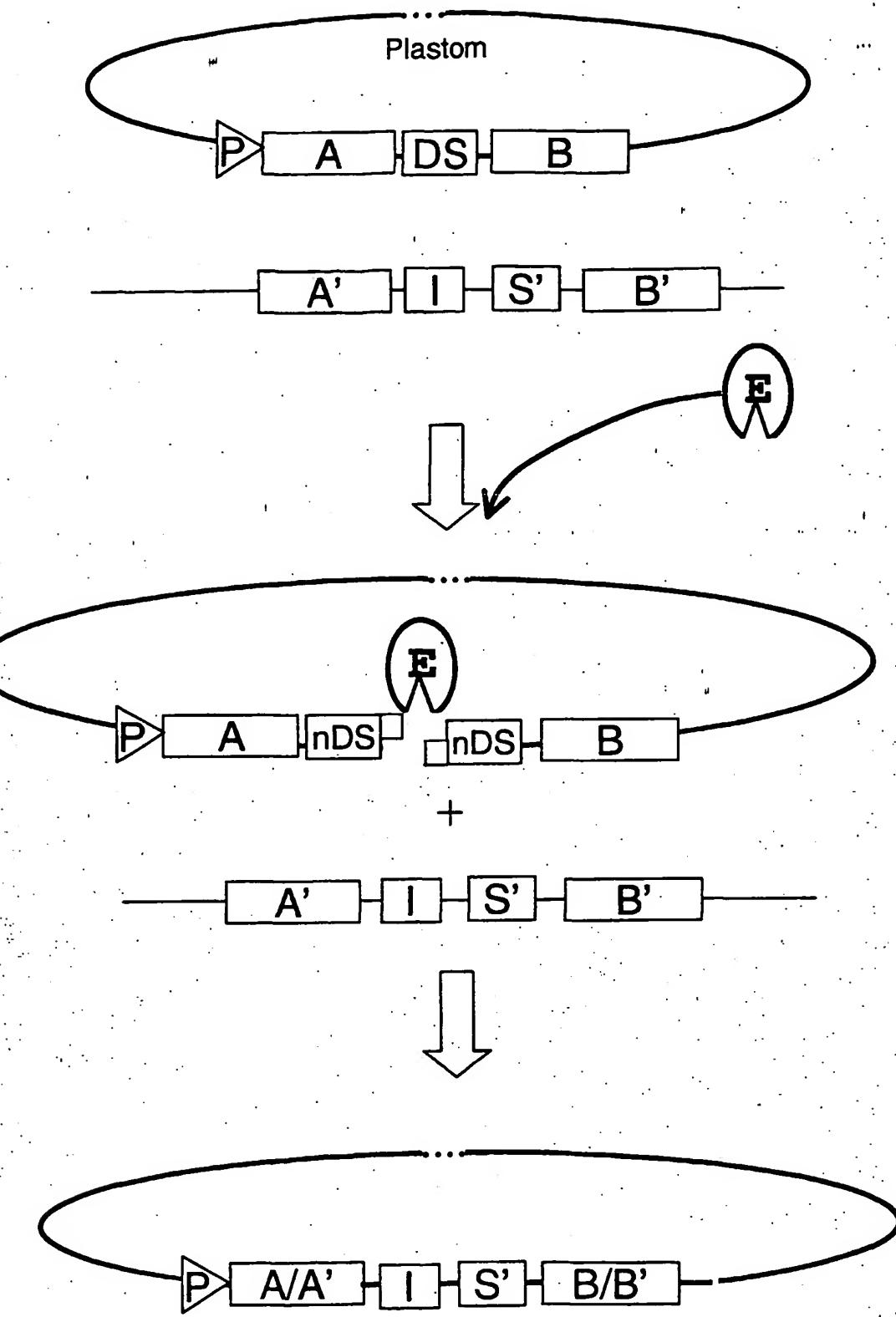


Fig. 6C

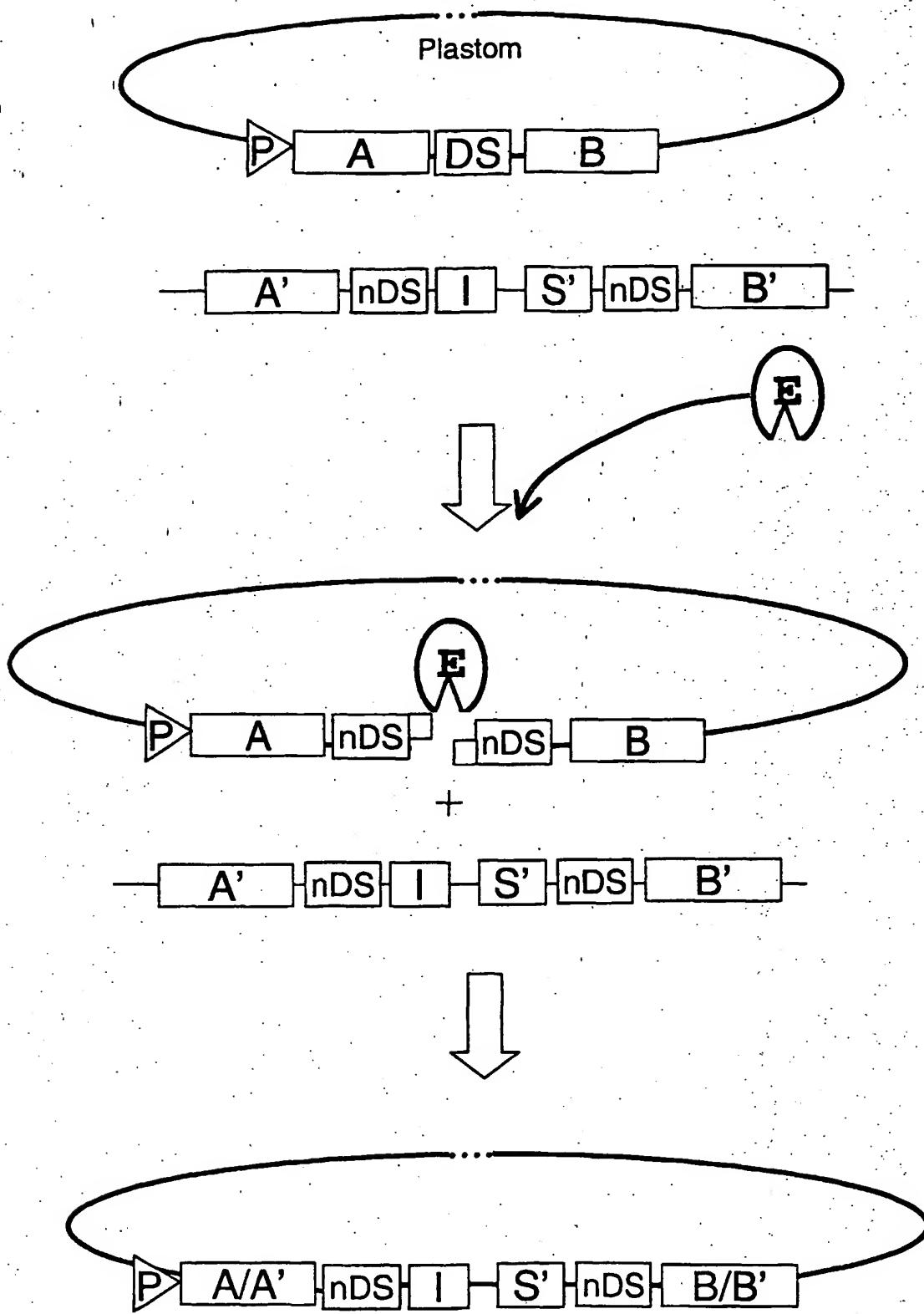


Fig. 6D

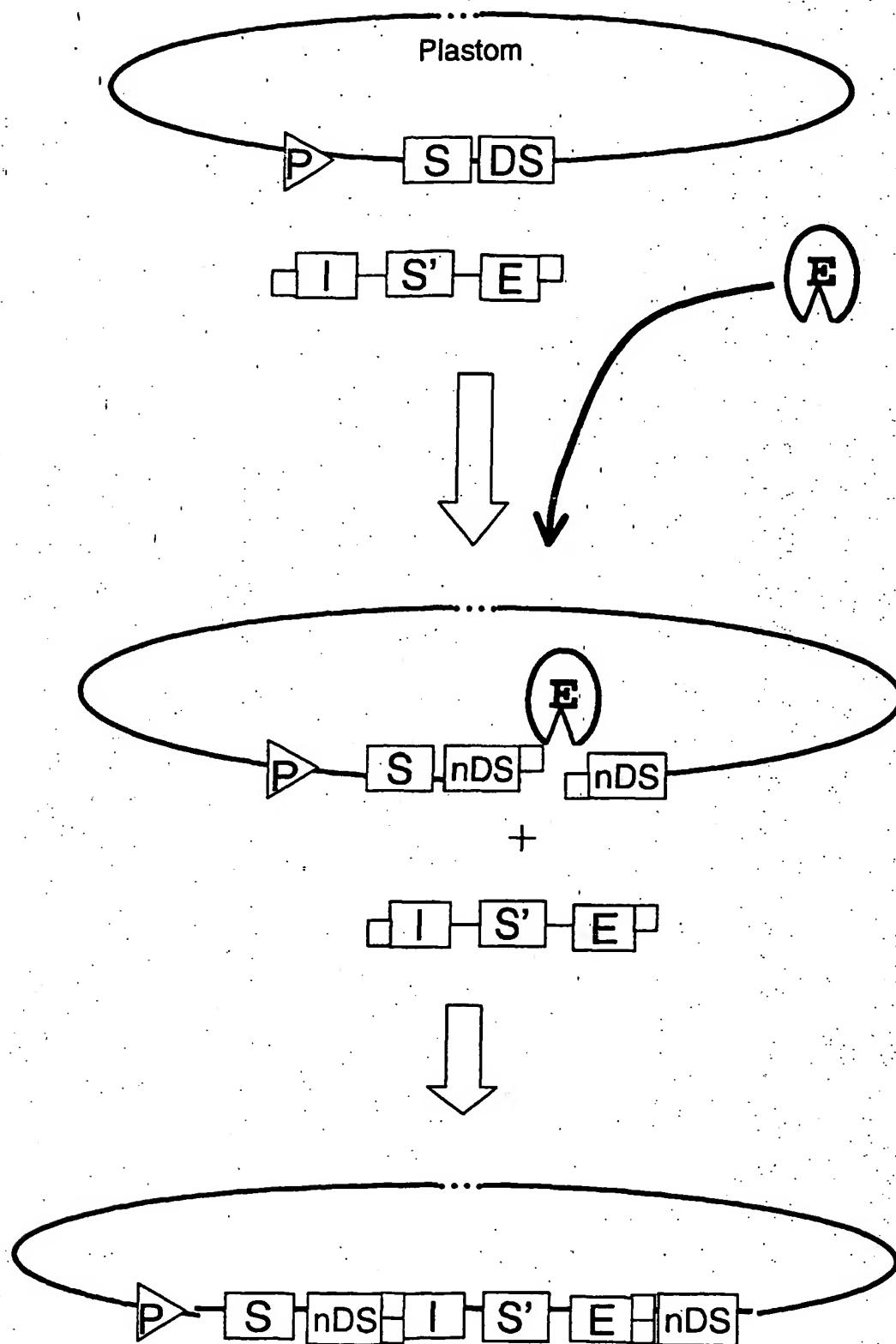


Fig. 6E

23/32

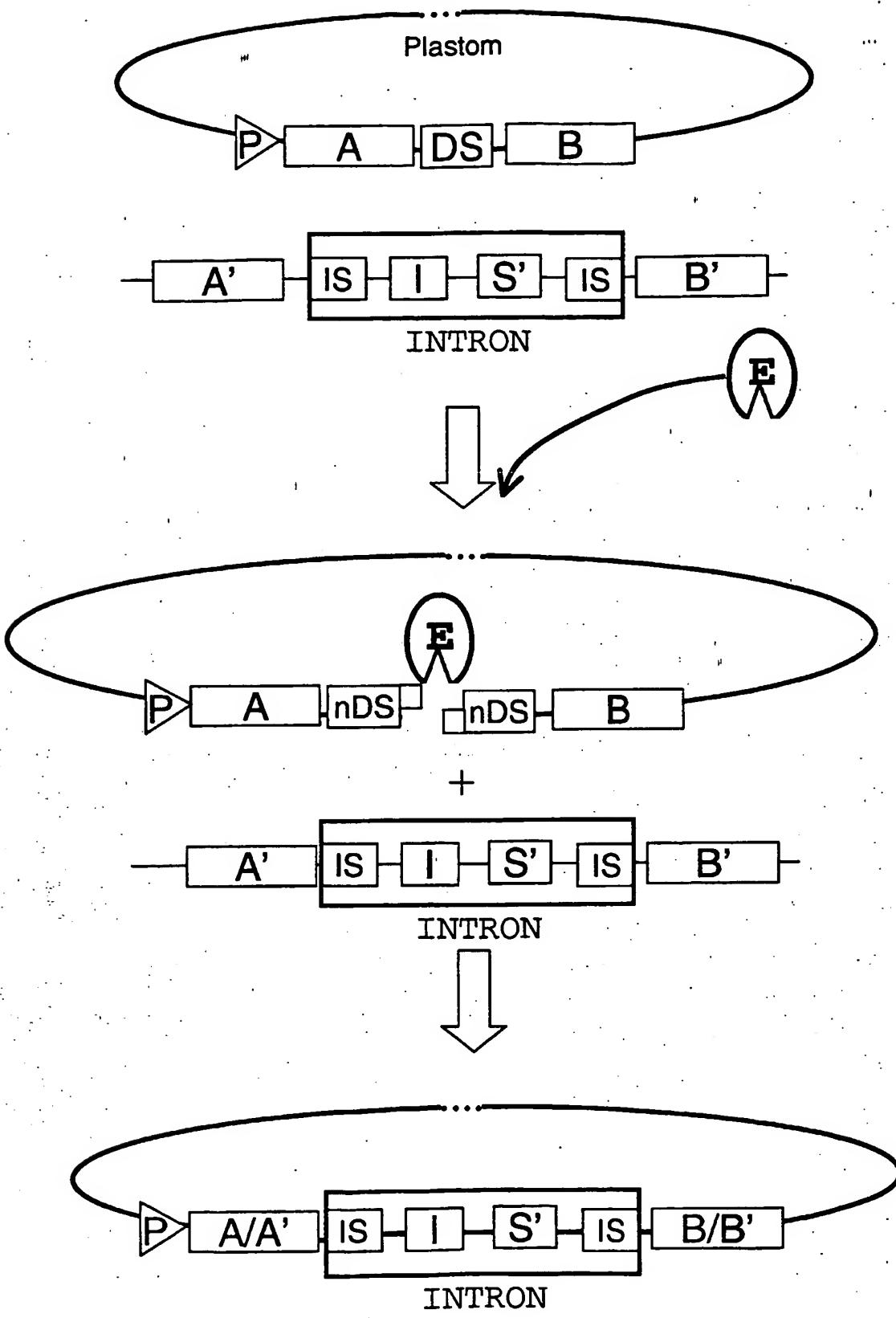


Fig. 7A

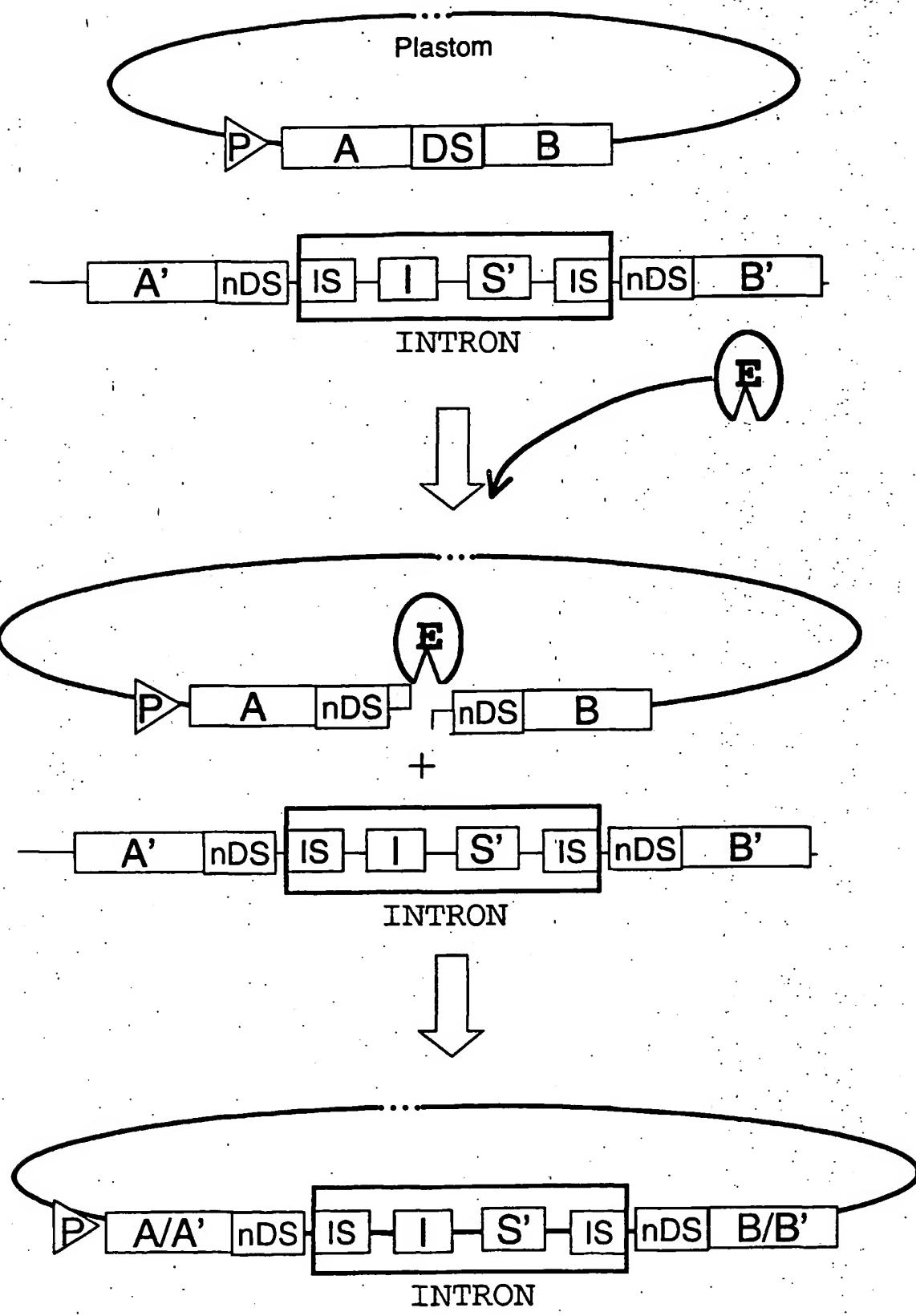


Fig. 7B

25/32

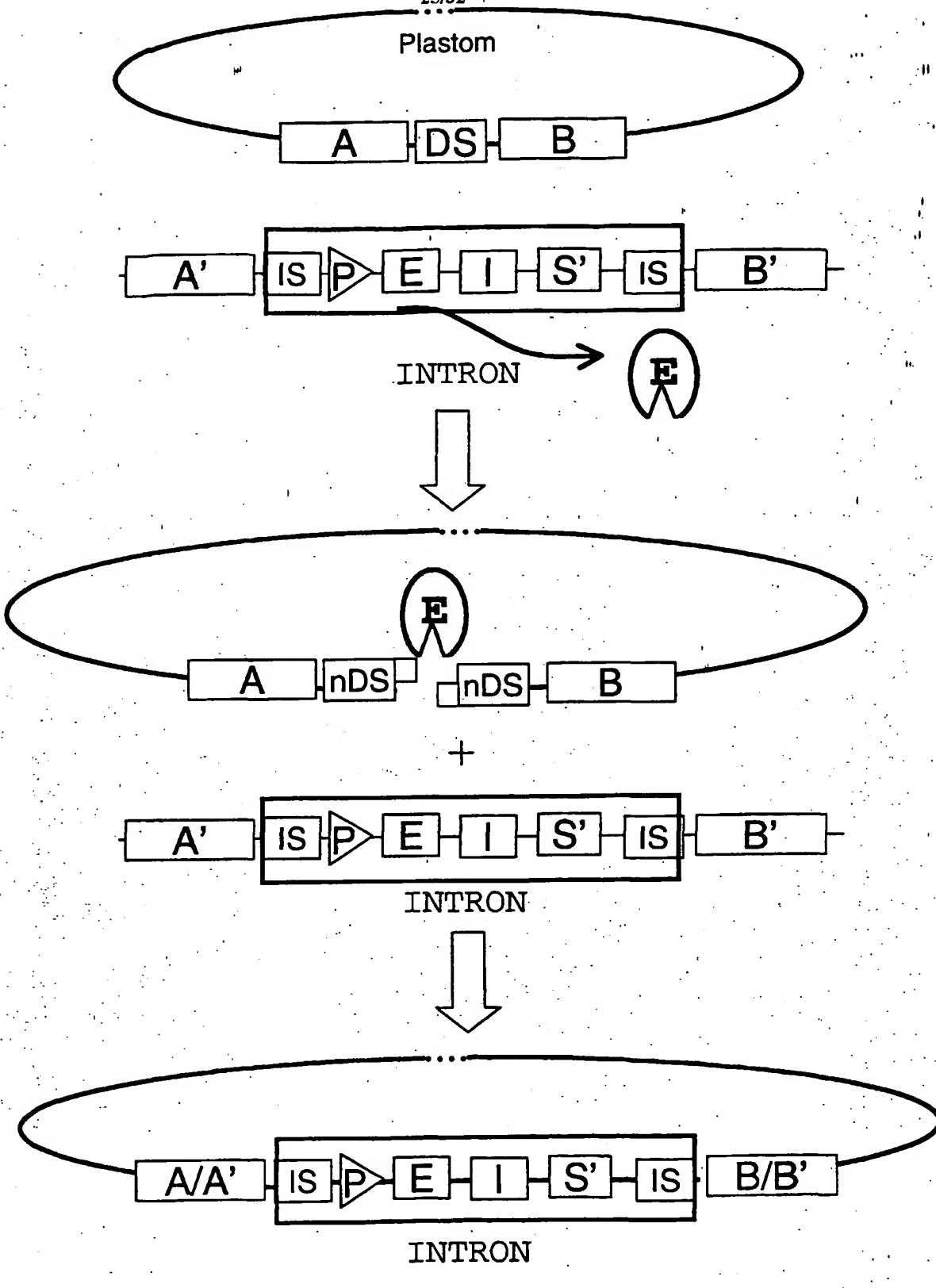


Fig. 7C

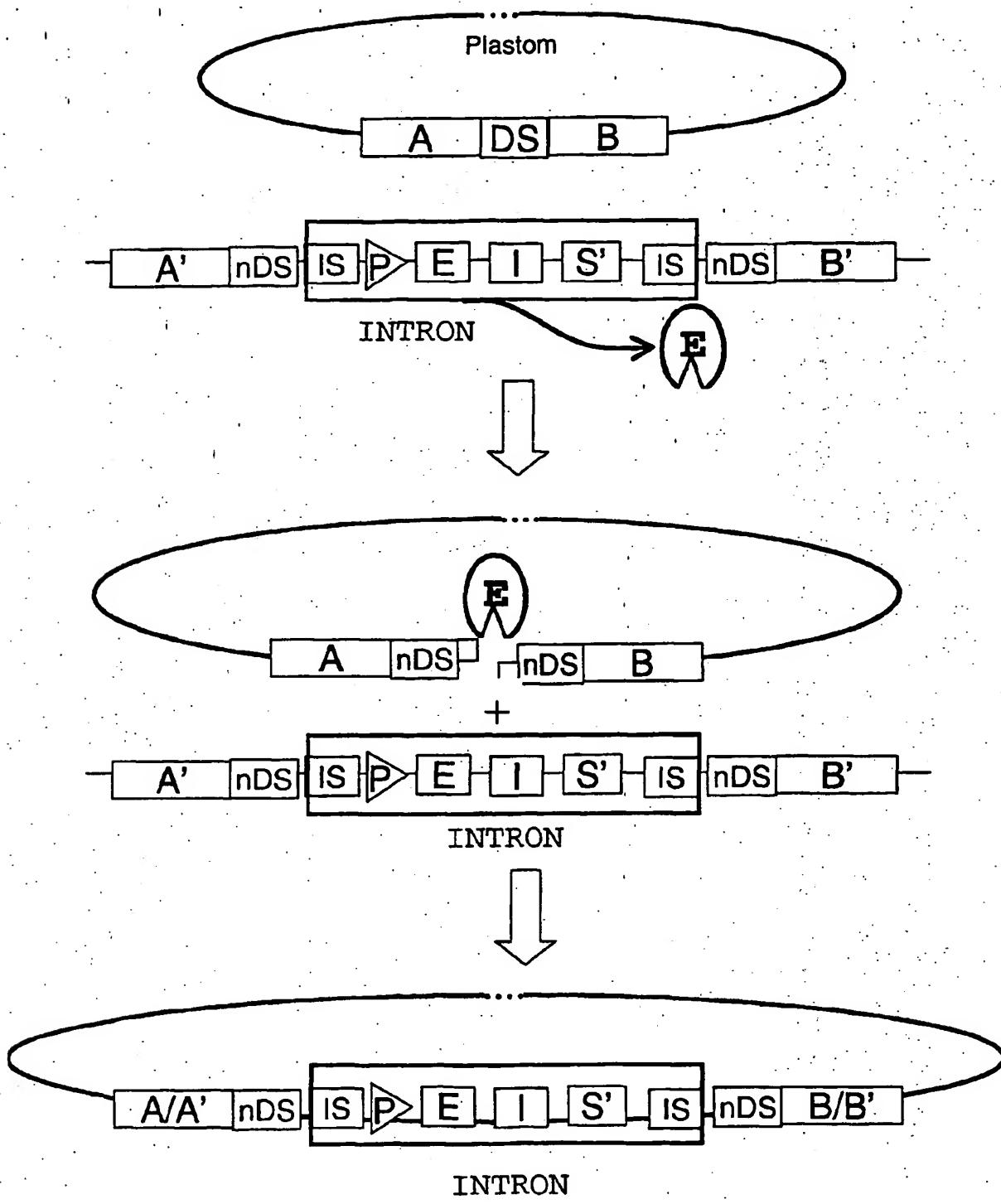


Fig. 7D

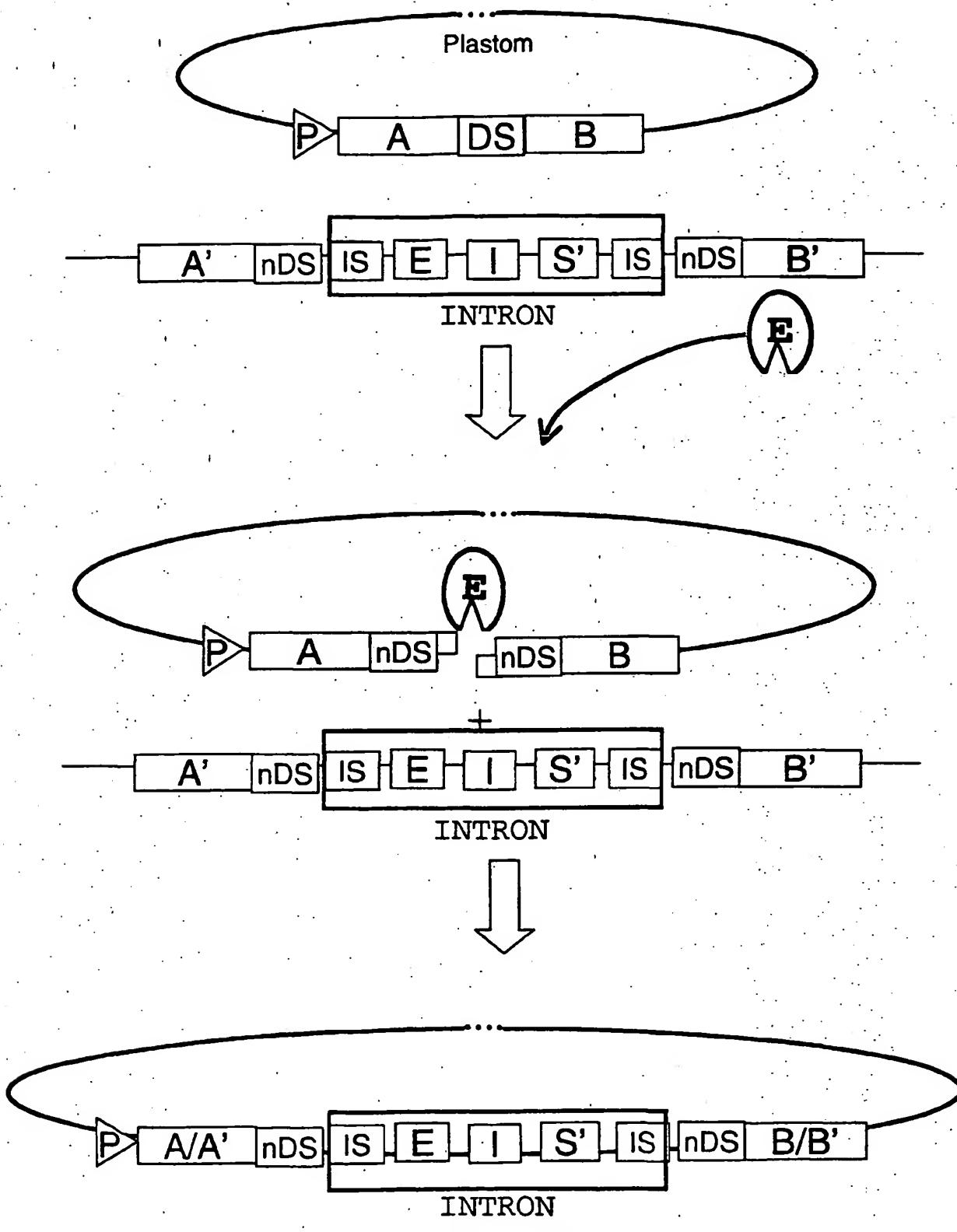


Fig. 7E

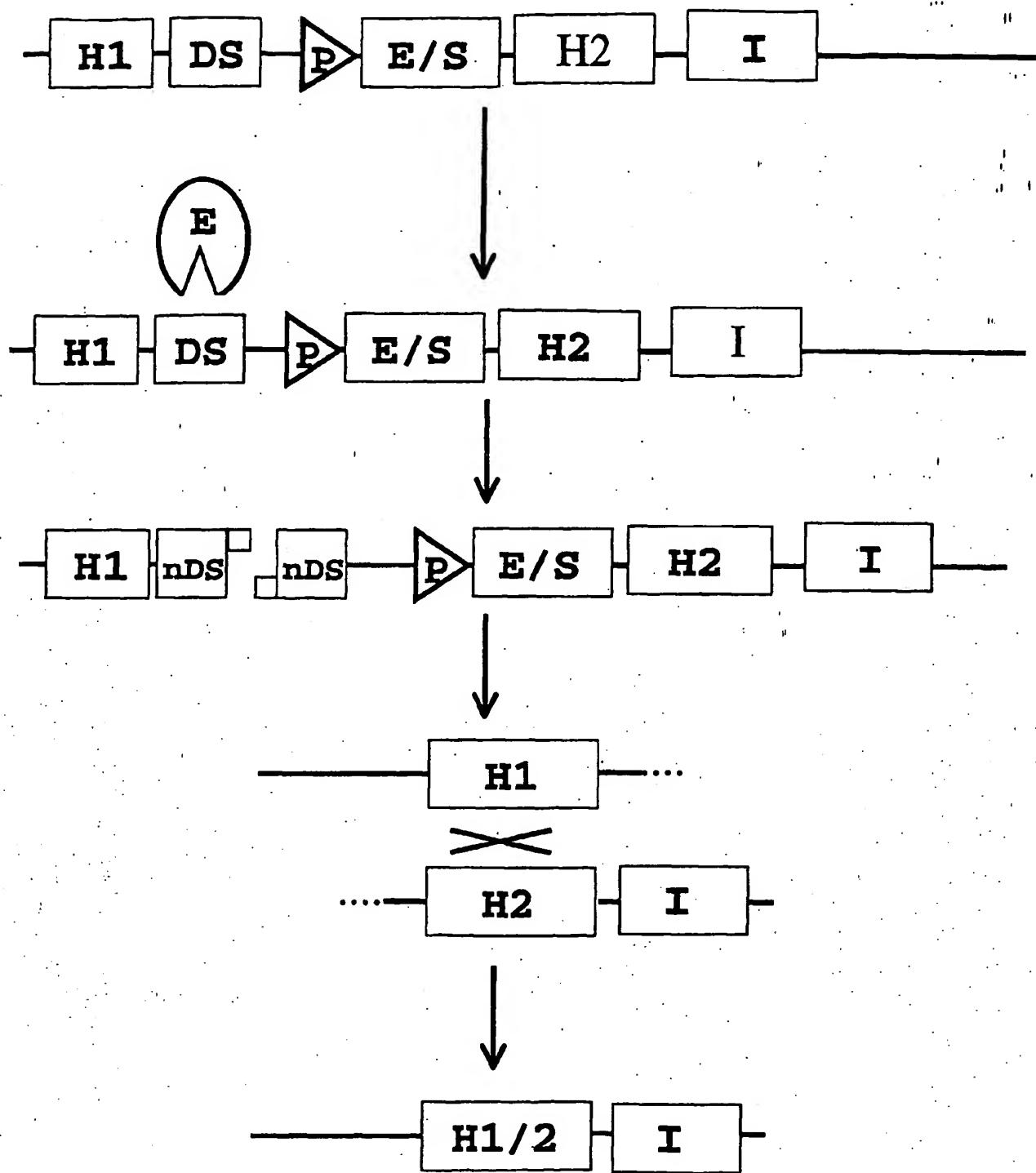
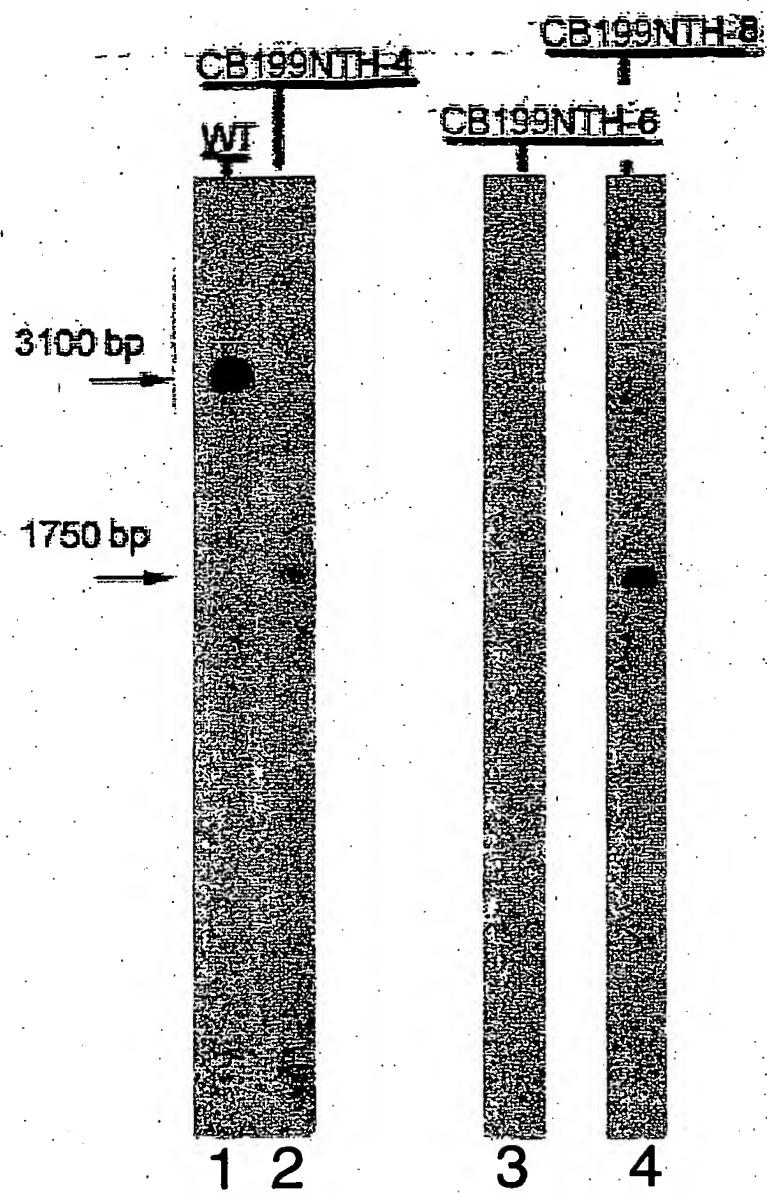
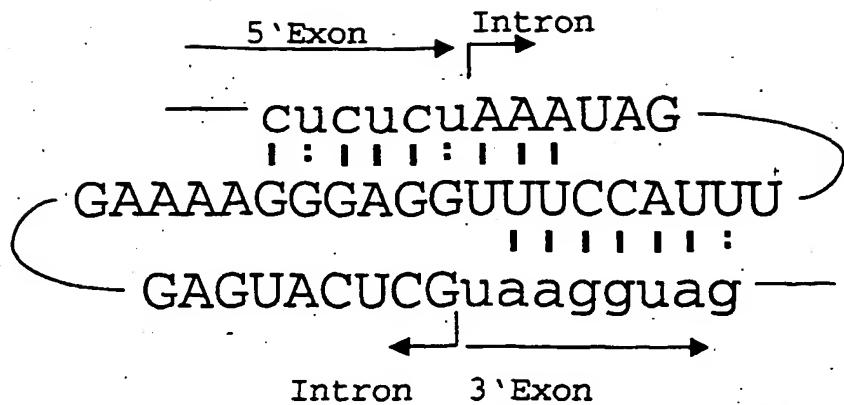


Fig. 8



**Fig. 9**

A



B

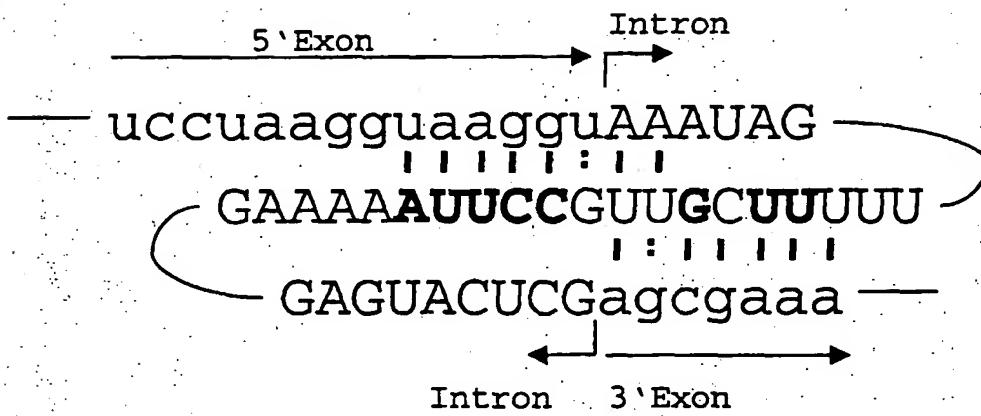


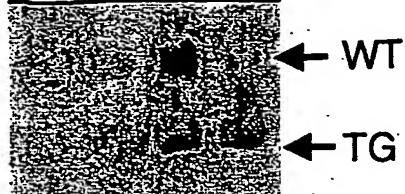
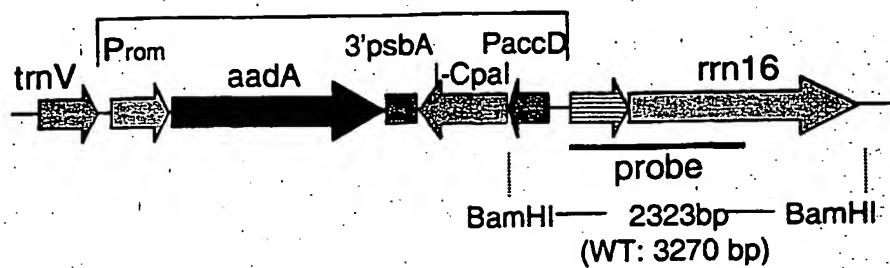
Fig. 10

**A**

CB255+435NTH

16b 16c 19 20

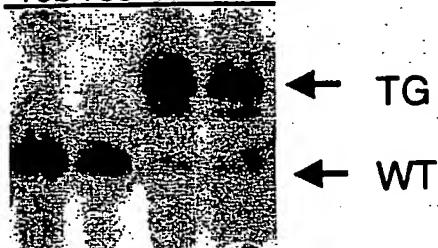
BamHI

**B****C**

CB255+435NTH

16b 16c 19 20

HindIII

**D**

23SrDNA Intron

probe

HindIII — 1511bp — HindIII  
(WT: 1098 bp)

Fig. 11

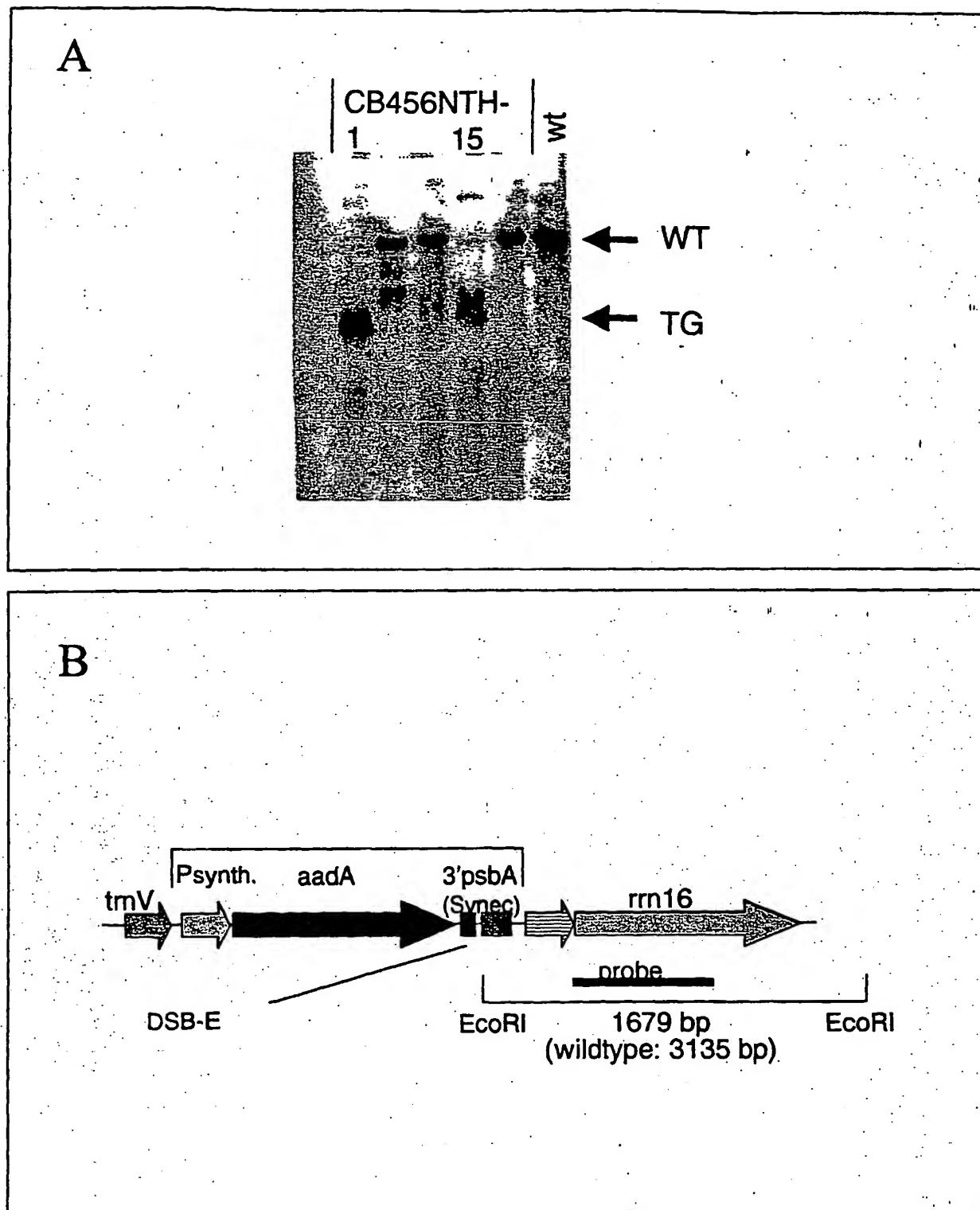


Fig. 12

## SEQUENZPROTOKOLL

<110> SunGene GmbH & Co.KGaA  
<120> Verfahren zur Transformation von pflanzlichen Plastiden  
<130> AE20010854 / PF 0081700028  
<140>  
<141>  
<160> 84  
<170> PatentIn Ver. 2.1  
<210> 1  
<211> 4363  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz: plastid transformation plasmid pCB42-94  
<220>  
<221> misc\_feature  
<222> (55)..(1405)  
<223> right targeting region  
<220>  
<221> mutation  
<222> (346)  
<223> mutation causing streptomycin resistance  
<220>  
<221> mutation  
<222> (68)  
<223> mutation causing spectinomycin resistance  
<220>  
<221> misc\_feature  
<222> (1404)..(1511)  
<223> multiple cloning site  
<220>  
<221> misc\_feature  
<222> (2629)..(3417)  
<223> Ampicillin resistance  
<400> 1  
gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagtgtatga 60  
cttgacggca tcctcacctt cctccggctt atcaccggca gtctgttcag gttccaac 120  
tcaacatgg caactaaaca cgagggttgc gctcggtcg ggacttaacc caacaccta 180  
cggcacgagc tgacgacagc catgaccac ctgtgtccgc gttcccgaag gcacccctt 240  
cttcaagag gattcgccgc atgtcaagcc ctggtaaggt tcttcgctt gcatcgaatt 300  
aaaccacatg ctccaccgct tgtgcgggccc cccgtcaatt ccttttagtt tcattttgc 360  
gaacgtactc cccaggcggg atacttaacg cgtagctac agcaactgcac gggtagatac 420  
gcacagcgcc tagtatccat cgtttacggc taggactact ggggtatcta atccccattcg 480  
ctccccctagc ttctcgctct cagtgtcagt gtcggcccaag cagagtgcctt tcggcggtgg 540  
tgttcttcc gatctctacg catttcaccg ctccaccgga aattccctct gcccctaccg 600  
tactccagct tggtagttc caccgcctgt ccagggttga gccctggat ttgacggccg 660  
actaaaaaag ccacctacag acgctttacg ccaatcatt ccggataacg cttgcacatct 720  
ctgtattacc gcccgtgctg gcacagagtt agccgatgtt tattccccag ataccgtcat 780  
tgctttcttcc cccggaaaag aagttcacga cccgtggggc ttctacaccttcc acgcggcatt 840  
gctccgtcag gtttcgccc attgcccact gtcgcctccc gtaggagtct 900

gggccgtgtc tcagtcccag tggctgat cattctcg gaccagctac tgatcatcg 960  
 cttggtaagc tattgcctca ccaactagct aatcagacgc gagccctcc tcggcggat 1020  
 tcctccttt gtcctcagc ctacgggta ttagcagccg tttcagctg ttgtcccc 1080  
 cccaaggca ggttttaacg cgtaactcac cctccgcca ctggaaacac cacttccgt 1140  
 ccgacttgc tggtaagc atgcccag cgtaactcac gagccaggat cgaactctcc 1200  
 atgagattca tagttcatt actatactc tccttgc tagacaagc ggattcggaa 1260  
 ttgtcttca ttccaaggca taacttgtat ccatgcgtt catattcgcc cggagttcgc 1320  
 tcccagaaat atagccatcc ctgcccctc acgtcaatcc cacgagcctc ttatccatc 1380  
 tcattgaacg acggcgaaaa agcgagctcc accgcggg tggccgtct agaacttagt 1440  
 gatccccgg gctcaggaa ttcgatatac agcttatacg taccgtgc acgtccggg 1500  
 ggccgggtac caaatccaa tagaaaaact cacattggc ttaggataa tcaggctcg 1560  
 actgtacttccaccgt caagtgaca ctctaccgt gagttatatac cttccccgc 1620  
 cccatcgaga aatagaactg actaatacta agtcaaagg tcgagaaact caacgccact 1680  
 attctgaac aacttggagc cgggccttct ttgcacta ttacggatataa 1740  
 gtcaaaatcg gattcaattt tcaactgcctt ctatcgaaa taggattgac taccgattcc 1800  
 gaaggaactg gagttacatc tctttccat tcaagagttc ttatgcgtt ccacgccc 1860  
 tttagacccc gaaaaatgga caaatccctt ttcttaggaa cacatacaag attcgtact 1920  
 aaaaaaaaaa taatggtaac cttaccatc actacttcat ttatgaattt catagtaata 1980  
 gaaatacatg tcctaccgag acagaatttgc ttccttgcc tagcaggcaa 2040  
 agatttaccc cctggaaag gatgattcat tcggatcgac atgagagtcc aactacattt 2100  
 ccagaatcca tggatatac ttgaaagagg ttgacccctt tgcttctctc atggtacact 2160  
 ccttccccgg ccgagccctt tttctctcg gtccacagag acaaataatgtt ggactgggtc 2220  
 caacaattca tcagactcac taagtcgggta tcaactaacta atactaatctt aatataatag 2280  
 tctaataatc ttaatataat agaaaaactt aatataatag aaaagaactg tctttctgt 2340  
 atacttccc cgggtccgtt gctaccgagg gcctcgat acgccttattt ttataggtt 2400  
 atgtcatgt aataatgggt tcttagacgt caggtggcac tttcgggaa aatgtcgccg 2460  
 gaacccttat ttgttattt ttctaaatac attcaataat gtatccgtt atgagacaat 2520  
 aaccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacattcc 2580  
 gtgtccccc tattccctt ttgcggcat ttgccttcc tgttttgtt caccagaaaa 2640  
 cgctgtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgac 2700  
 tggatctcaa cagcgtaag atccttgaga gtttcgccc cgaagaacgt tttccatgt 2760  
 tgagcactt taaagttctg ctatgtggcg cggattttatc ccgtatttgc gcccggcaag 2820  
 agcaactcgg tcgcccata cactattctc agaatgactt gttgagttc tcaccagtc 2880  
 cagaaaagca ttttacggat ggcatgacag taagagaatt atgcagtgtt gccataacca 2940  
 tgagtataa cactgcggcc aacttacttc tgacaacgtt cggaggaccg aaggagctaa 3000  
 cccctttttt gcacaacatg gggatcatg taactcgcc ttagtgcgtt gaaaccggagc 3060  
 tgaatgaagc cataccaaac gacgagcgtt acaccacgt gcctgttagca atggcaacaa 3120  
 cgttgcgcaa actattaact ggcgaactac ttactctgtt ttccggcaa caattaatag 3180  
 actggatgga ggcggataaa gttgcaggac cacttctcg ctggccctt ccggctggct 3240  
 ggtttattgc tgataaatct ggagccgggtg agcgtgggtc tcgcgttattc attgcagcac 3300  
 tggggccaga tggtaagccc tccgtatcg tagttatcta cacgacgggg agtcaggcaa 3360  
 ctatggatga acgaaataga cagatcgctg agataggtgc ctactgtt aagcattgtt 3420  
 aactgtcaga ccaagtttac tcatatatac tttagattgtt tttaaaactt catttttaat 3480  
 ttaaaaaggat cttagtgaag atccttttg ataatctcat gacaaaatc ctttaacgtt 3540  
 agtttcgtt ccactgagcg ttagaccccg tagaaaaagat caaaggatct ttttgcgtt 3600  
 cttttttctt ggcgttaatc tgctgcgttgc aaacaaaaaa accaccgtt ccagcggtgg 3660  
 ttgtttgc ggtatcaagag ctaccaactc ttttccgaa ggtactggc ttcaagcag 3720  
 cgcagataacc aaataactgtc ttcttagtgc agccgttagtt aggccaccac ttcaagaact 3780  
 ctgttagcacc gcctacatac ctgcgttgc taatcctgtt accagtggct gtcggcgtt 3840  
 gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 3900  
 ggtcgggctg aacgggggggt tcgtgcacac agcccagtt ggagcgaacg acctacacccg 3960  
 aactgagata cctacagcgt gagctatgag aaagcggccac gttccggaa gggagaaaagg 4020  
 cggacaggta tccggtaagc ggcagggtcg gaacaggaga ggcacgagg gagttccag 4080  
 ggggaaacgc ctggtatctt tatagtcctg tcgggttgc ccacctctga ttgcgttgc 4140  
 gattttgcgtt atgctgtca gggggcggaa gcctatggaa aaacgccagc aacgcggcct 4200  
 ttttacgggtt cttggccctt tgctggccat gttcttccat gcttccatc 4260

ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataaccgt cggccgagcc 4320  
 gaacgaccga ggcgcagcgag tcagtgagcg aggaagcgga aga 4363

<210> 2

<211> 1359

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Insert of vector pCB199-3

<400> 2

gagctctgat cacggaagat agcttggca aaaaaagcaa aaagcattta ctttgattga 60  
 gatgttaatt gtgttggcaa ttatcgttat tttatatttgc caaatggat 120  
 actagagctt cgggtgccag ggcgtccct tgggtcccc gggcgctac tcgacgctac 180  
 cttaaagagag tcaagcttc atattaccct gttatcccta gcgtactcga gaaaaaaaaa 240  
 aaggagcaat agcaccctct tgatagaaca agaaaaatgtat tattgtctt ttctttcaa 300  
 aaccccttat agactaggcc aggaatttac tgcaattttc tgccaactac ctttagtgatc 360  
 tcgccttca cgtatgtggac aaattttcc aactgtatctg cgccgcggc caagcgatct 420  
 tcttcttgatc caagataagc ctgtcttagt tcaagatgtatc cgggtgtata ctggggccggc 480  
 aggccgttcca ttgcccagtc ggcagcgaca tccctcgcc cgattttgcg ggttactcg 540  
 ctgtacccaa tgcgggacaa cgtaaagact acatttcgct catcccgccg ccagtcgggc 600  
 ggcgagttcc atagcgatc ggtttcattt agccctcaaa atagatcttgc ttccaggaaacc 660  
 ggtatcaaaga gttccctccgc cgctgaccc accaaggcaaa cgctatgttc tcttgctttt 720  
 gtcagcaaga tagccagatc aatgtcgatc gtggctggct cgaagatacc tgcaagaatg 780  
 tcatttgcgtt gccattttcc aaatttgcgtt tcgcgtttag ctggataacg ccacggaaatg 840  
 atgtcgatcg gcaacaacaat ggtgacttct acagcgccgaa gaatctcgat ctctccaggg 900  
 gaagccgaag ttccaaaag gtcgttgatc aaagctcgcc gctgttttc atcaagcctt 960  
 acggtcaccg taaccagcaaa atcaatatca ctgtgtggct tcaggccgatc atccactcg 1020  
 gagccgtaca aatgtacggc cagcaacgtc ggttgcgat ggcgtcgat gacgccaact 1080  
 acctctgata gttgagttga tacttcggcg ataaccgctt cacgagccat ggctgtttgt 1140  
 ggtgtcatgg ctgtttgtgg tgcgtatc ccctccctac aactagatcc tcgcccggag 1200  
 ttccgtccca gaaatatacg catccctgat ccctcacgtc aatccacga gccttttac 1260  
 cattctcattt gaaacgacggc gtcgaggggg ggcccggtac gtcgacgaaat ttccatttcc 1320  
 gaagttccca ttctcaagaa agtatacgaa cttcgttacc 1359

<210> 3

<211> 1070

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insert of vector pCB401-20

<400> 3

gagctctgat cacggaagat agcttggca aaaaaagcaa aaagcattta ctttgattga 60  
 gatgttaatt gtgttggcaa ttatcgttat tttatatttgc caaatggat 120  
 actagagctt cgggtgccag ggcgtccct tgggtcccc gggcgctac tcgacgctac 180  
 cttaaagagag tcaagcttc atattaccct gttatcccta gcgtactcga gctgtgtt 240  
 ttccgtccact accttagtgc ttcgtccctt cacgtatgtt gacaaattttt ccaactgtatc 300  
 tcgcgtcgat gccaagcgat ttcttcttgc tccaaagatc gctgtcttag ctcaagtatc 360  
 gacgggtgtatc tactggccgc gcaaggcgatc cattggccatc tcggcagcgatc catccctcg 420  
 cgccgttttgc cccgttactgt cgctgtatc aatgcgggac aacgtaaatcgat ctacatttgc 480  
 ctcatcgatc gcccacgtcgat gccggcgtt ccatagcgat aaggtttcat ttagcgccctc 540  
 aaatagatcc tgcgtatc ccggatcataa gagttccctcc gccgtggac ctaccaaggc 600  
 aacgctatgt ttcgttgcgtt ttgtcagcaaa gatagccaga tcaatgtcgat tcgtggctgg 660  
 ctcgaagata cctgcgtatc tgcgtatgttgc tgcgtatgttgc ccaatttgc gttcgccgtt 720  
 agctggatataa cggccacggaa tgcgtatgtcgat gtcgacaaaca atgggtactt ctacagcgatc 780  
 gagaatctcg ctctctccat gggaaagccgat ggtttccaaagggatcgttgc tcaaaatgtcg 840

ccgcgttgtt tcatacgcc ttacggcacg cgtaccacgc aaatcaatat cactgtgtgg 900  
 cttcaggccg ccatccactg cggagccgta caaatgtacg gccagcaacg tcggttcgag 960  
 atggcgctcg atgacgccaa ctacctctga tagttgagg tatacttcgg cgataaccgc 1020  
 ttcacgagcc atggttccct ccctacaacg tcgagggggg gccccgttacc 1070

&lt;210&gt; 4

&lt;211&gt; 804

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; Beschreibung der künstlichen Sequenz: insert of vector pCB289-13

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (154)...(642)

&lt;223&gt; sequence coding for I-PpoI homing-endonuclease

&lt;400&gt; 4

ggtaccgggc ccccccctcga cgccgtcggtt caatgagaat ggataagagg ctcgtggat 60  
 tgacgtgagg gggcaggat ggcttatattt ctgggagcga actccggcg aggatctagt 120  
 tgttagggagg gattcatgac accacaaaaca gcc atg gcg ctc acc aat gct caa 174  
 Met Ala Leu Thr Asn Ala Gln

1

5

atc ttg gct gtg att gac agt tgg gaa gaa aca gtc ggt cag ttt cca 222  
 Ile Leu Ala Val Ile Asp Ser Trp Glu Glu Thr Val Gly Gln Phe Pro  
 10 15 20

gtg ata acg cac cat gta cca tta ggt ggc ggt ctg caa gga acg ctc 270  
 Val Ile Thr His His Val Pro Leu Gly Gly Leu Gln Gly Thr Leu  
 25 30 35

cat tgt tac gag atc ccc cta gca gct cct tat ggg gtt ggc ttt gct 318  
 His Cys Tyr Glu Ile Pro Leu Ala Ala Pro Tyr Gly Val Gly Phe Ala  
 40 45 50 55

aag aat ggg cct acc cgc tgg caa tac aaa cgg aca atc aat caa gtc 366  
 Lys Asn Gly Pro Thr Arg Trp Gln Tyr Lys Arg Thr Ile Asn Gln Val  
 60 65 70

gtc cac aga tgg ggg tcc cac aca gtc cct ttt cta tta gaa ccg gat 414  
 Val His Arg Trp Gly Ser His Thr Val Pro Phe Leu Leu Glu Pro Asp  
 75 80 85

aac atc aac ggc aaa acc tgc aca gca tcg cac cta tgt cat aat act 462  
 Asn Ile Asn Gly Lys Thr Cys Thr Ala Ser His Leu Cys His Asn Thr  
 90 95 100

cga tgc cac aat ccc ttg cac tgc tgg gag tca cta gac gac aac 510  
 Arg Cys His Asn Pro Leu His Leu Cys Trp Glu Ser Leu Asp Asp Asn  
 105 110 115

aaa ggc aga aac tgg tgc ccc ggt ccc aac ggg gga tgt gtc cat gcg 558  
 Lys Gly Arg Asn Trp Cys Pro Gly Pro Asn Gly Gly Cys Val His Ala  
 120 125 130 135

gtg gtt tgt tta agg cag ggt ccg ttg tac ggc cca ggg gcg act gtg 606  
 Val Val Cys Leu Arg Gln Gly Pro Leu Tyr Gly Pro Gly Ala Thr Val  
 140 145 150

gca ggt cct caa caa agg ggc agt cac ttt gtg gta taactgcaga 652  
 Ala Gly Pro Gln Gln Arg Gly Ser His Phe Val Val  
 155 160

agcttcaatt gcatgctcta gatgatcaaa gaattcctgg cctagtctat aggaggttt 712  
 gaaaagaaaag gagcaataat catttcttg ttcttatcaag agggtgctat tgctccttc 772  
 ttttttctc gagaggggta cgtaccgagc tc 804

&lt;210&gt; 5

&lt;211&gt; 163

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;223&gt; Beschreibung der künstlichen Sequenz: insert of vector pCB289-13

&lt;400&gt; 5

Met Ala Leu Thr Asn Ala Gln Ile Leu Ala Val Ile Asp Ser Trp Glu			
1	5	10	15

Glu Thr Val Gly Gln Phe Pro Val Ile Thr His His Val Pro Leu Gly			
20	25	30	

Gly Gly Leu Gln Gly Thr Leu His Cys Tyr Glu Ile Pro Leu Ala Ala			
35	40	45	

Pro Tyr Gly Val Gly Phe Ala Lys Asn Gly Pro Thr Arg Trp Gln Tyr			
50	55	60	

Lys Arg Thr Ile Asn Gln Val Val His Arg Trp Gly Ser His Thr Val			
65	70	75	80

Pro Phe Leu Leu Glu Pro Asp Asn Ile Asn Gly Lys Thr Cys Thr Ala			
85	90	95	

Ser His Leu Cys His Asn Thr Arg Cys His Asn Pro Leu His Leu Cys			
100	105	110	

Trp Glu Ser Leu Asp Asp Asn Lys Gly Arg Asn Trp Cys Pro Gly Pro			
115	120	125	

Asn Gly Gly Cys Val His Ala Val Val Cys Leu Arg Gln Gly Pro Leu			
130	135	140	

Tyr Gly Pro Gly Ala Thr Val Ala Gly Pro Gln Gln Arg Gly Ser His			
145	150	155	160

Phe Val Val

&lt;210&gt; 6

&lt;211&gt; 1150

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
 XbaI / Bgl II fragment for construction of vector pCB304-25

&lt;400&gt; 6

ctcgagtatt cggctcaatc ctttagtaa aagattgggc cgagttaat tgcaattcaa	60
ttaagagaac gaaggataat tacttgatt ctttctcctt atccttcttt atttcctgt	120
aatttatctg ctaatgtcta ctgaattcaa gcttggatcc gggccgcct agcttgggtc	180
ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg	240
cgataccgta aagcacgagg aagcgtcag cccattcgcc gccaagctct tcagcaatat	300
cacggtagc caacgctatg tcctgatagc ggtccgcac acccagccgg ccacagtcta	360
tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca tcgccatggg	420

tcacgaacgag	atccctcgccg	tcgggcatgc	gcgcctttag	cctggcgaac	attcggctg	480
gcgcgagccc	ctgatgtct	tcgtccagat	cattctgatc	gacaagaccg	gtttccatcc	540
gagtacgtgc	tcgctcgatg	cgatgttgc	cttggggc	aatgggcag	gtagccggat	600
caaggctatg	caggccgcgc	attgcattcag	ccatgtatgg	tactttctcg	gcaggagcaa	660
ggtagatgatg	caggagatcc	tgccccggca	cttcgccccaa	tagcagccag	tcccttcccg	720
cttcagtgac	aacgtcgagc	acagctgcgc	aaggaacgc	cgtcgtggcc	agccacgata	780
gcccgcgtgc	ctcgctctgc	agttcattca	gggcacccgg	caggtcggtc	ttgacaaaaaa	840
gaaccgggcg	ccctcgct	gacagccg	acacggccgc	atcagagcag	ccgattgtct	900
gttgtgcca	gtcatagccg	aatagccct	ccacccaagc	ggccggagaa	cctgcgtgca	960
atccatcttgc	ttcaatccaa	gctcccatgg	ctgtttgtgg	tgtcatgaat	ccctccctac	1020
aactagatcc	tatatactat	agagataggt	ggataaaat	ttttctttag	taagacccca	1080
tcgctaata	taatttatct	aacatattaa	ttaatattta	atatataaaat	atataatagag	1140
tcgagatcta						1150

<210> 7

<211> 454

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insert of vector.pCB220-17

<400> 7

gaagacgatc ctaaatagca atatttacct ttgggaccaa aagttatcg gcatgcacct 60  
ggtagctagt cttaaacca atagattgca tcggttaaa aggcaagacc gtcaaattgc 120  
gggaaagggg tcaacagccg ttcatgtacca agtctcaggg gaaactttga gatggccttg 180  
caaagggtat ggtataataa tgacggacat ggtcctaacc acgcagccaa gtcttaagtc 240  
aacagatctt ctgttgatat ggatgcagtt cacagactaa atgtcggtcg gggaaagatgt 300  
attcttctca taagatatac tcggactctc ccttaatggg agctagcggg tgaagtgtatg 360  
caacacttggc gcccgtggc actaatttgt atgcggaaatgt atattgatta gttttggagt 420  
actcgaaggt agcgaaattc cttgtcaagt ctcc . . . . . 454

<210> 8

<211> 498

<212> DNA

### <213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insert of  
vector PCB270-1

<400> 8

gcgcgcgtaa tacgactcac tatagggcga attggggcg actgtttacc aaaaacacag 60  
gtctccgcaa agtcgtaaga ccatgtatgg gggctgacgc ctgcccagtg cccgaaggtc 120  
aaggaagttg gtgacctgat gacaggggag cccgcgaccg aagccccgt gaacggcggc 180  
cgtaactata acggtcctaa gtctcgaat cgaattcaact agtgatagga agaccttgc 240  
gggtaagttc cgacccgcac gaaaggcgta acgatctggg cactgtctcg gagagaggct 300  
cggtgaaata gacatgtctg tgaagatgcg gactacctgc acctggacag agagacccca 360  
attcctggcc tagtctatag gaggtttga aaagaaaagga gcaataatca ttttcttgtt 420  
ctatcaagag ggtgctattg ctcccttctt ttttctcga ccagcttttg tttccctttag 480  
tgagggttaa ttgcgcgc 498

<210> 9

<211> 789

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: lacZ gene with inserted intron to demonstrate splicing (from vector pCB315-1)

&lt;400&gt; 9

acaatttcca ttgcattc aggctgcga actgttggga agggcgatcg gtgcgggcct 60  
 ctgcgtat acgcccagctg gcgaaagggg gatgtgctgc aaggcgatta agttggtaa 120  
 cggccagggtt ttcccagtca cgacgttgc aaacgacggc cagtgagcgc gcgtataacg 180  
 actcactata gggcgaattt ggtaccggc cccccctcg a ggtcgacggt atcgataagc 240  
 ttgaccggac aaggaattt gctacccatcg agtactccaa aactaatcaa tatacttcg 300  
 catacaaattt agttcccgac ggctccagtg ttgcatcaact tcatccgcta gctccat 360  
 aggagaggc c gactatatac ttatgagaag aatacatctt ccccgaccga catttagtct 420  
 gtgaactgc a tccatataca cagaagatct gttgacttag gacttggctg cgtggtagg 480  
 accatgtccg tca gcttattt accataccct ttgcaaggcc atctcaaagt ttccccttag 540  
 acttggtaact gaacggctgt tgacccttt cccgcaattt gacggcttg cttttaaac 600  
 cgatgcaatc tattggttt aagacttagt accaggtgca tgcctgataa cttttggcc 660  
 caaagtaaa tattgctatt taggaccgtt agatccact agttctagag cggccgccc 720  
 cgcggggag ctccagctt tgcccttt agtgggggtt aattgcgcgc ttggcgtaat 780  
 catggcat 789

&lt;210&gt; 10

&lt;211&gt; 2566

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: L1.LtrB  
intron from vector pCB345-34.

&lt;400&gt; 10

gaattcgccc ttgaagact tgacatttga tatggtaaag tagggaggta ccgccttgg 60  
 cacattactg tgactgggtt gcaccaccct cttcgggaac cgtacgtacc cctctcgag 120  
 tatacggtc tggattttt cgtcgtaaa aattcaactt gtttatgaa tca cgtgacg 180  
 atgacaatga aagcatacaa caagatttt acgttggttc gctatcattt ccatttccca 240  
 ttttctttt ccttaagat ttttgcattt attgacatgg tgaatttcat aggaagtatt 300  
 ttcatcagat gttccacata attcacaaca ttagctttt aacctgtttt caagagtatt 360  
 cccggccatag ccatacaata caggagctt acttatctt tccgtaaattt gataagggg 420  
 ttacattca cttaaaattt caaaataacg ggcgtctta cttgtttt tctcatacgg 480  
 gatgccccac gaaccactt catctttaaa catggaaatg gttttgaaa gtgtccctt 540  
 atgttggag gctatcg ttagacagct gtattccata agataagcaa aataatttag 600  
 ctggtaaaa ttacttgcta gaccgttagt attacaatc cctcttaattt cagaattata 660  
 aattgtgatg atttctaagt ctgttgaacg aataagatatt tccctgtgaa ctggaaacca 720  
 tgagctatct ttcttttggaa tagtattttt cttgtcaaaa ataaatttgc gaaattttgtc 780  
 ttgaagagga ataaggagtt ctacactccc attgagtgtt ctcttttga ctttaccaga 840  
 tcgtttatc gttccacttc tccttactcg tatatcatat cccagaaaac gaggccgtt 900  
 actgctatgt gtgatgagtg tttttcttc actcaattcc atttttagt ttttatgat 960  
 aaaaatttt aatttttctt ttatccattt acagtcctt tgccttccctt taacagat 1020  
 aatgaagtgc tccgcataacc ggacgtatcc caatacttta tttgtctgtg aggtacaggg 1080  
 gagtgtgggt aatcttttac gttttctt gatattctaaa agaacttttag ctttttcttc 1140  
 accctccaaac ttcttgagac ggtgagaaat tcttttattt tcattgtgaa gttcccgata 1200  
 ttcagggtttt atttcttcg gacttctcg gtcaaaacttcc atttttagt gtaaaaacaaa 1260  
 cttatccaaat tcatgaagat agatgttggc caaaagagga gatagaattt cacccttgg 1320  
 tggccgtt gtaatgtt gatactgcca gttttccaga taacctgctt ttagaaatattt 1380  
 ataaatcaat tggctcattt tcatactttt gatttttaaga ttgatgagtc caatgagtg 1440  
 aacgtggctt atattatcg a g c g c t t t t t a t a t c t c c t c c a c c a t t c c 1500  
 gccaaactctt cttttgattt tttccaaagc cgtgtgacag cttcgttgg g t c t a a a a c c 1560  
 gtgagacaca ttttgcataa cccgttccata gatagattca agaatttattt tcacagcttc 1620  
 ttggatcaat ttatctgtgaa aagttggaaat tccctaaagggtt ctcatctttt tagaatttctt 1680  
 ttttgcataa tacatttttc gtacagggtt aggatgatgaa gttccgttctt taaaagatgg 1740  
 aataatctttt tttatctttt cttcaacttcc gccatccgct gatcatctt atattctttt 1800  
 tggaaagct ctttatttgg aatataaaattt ttgatcgc acgtaataaaa tatctggacg 1860  
 taaaagataa cgataaaagtc ttgtaaaaac ttgcgtctata ttttcttgg aattttactt 1920  
 gattcttctt aaaaatttgc tttttttttt cattttgggg ttttccctccca taatcaattt 1980

ttaattttac tacacaataa ctgtaccctt ttgccatgtaa aaggcggtta cccttcttag 2040  
actactacga gtactccgtaa cccttgcaag atttcaagc tctagtgtcta tagcctttt 2100  
cctcccttctt attaggcatt cttgtttagg gtatccccag ttatgtttaa gtctggtaa 2160  
attcagattc tcggcatcgc ttctgttcg ttcccataagg ttctcctaca gattgtacaa 2220  
atgtgtgtat aacagataag tccacaacca taacttaccc ttcttgta tagaggttc 2280  
agacactttc ctctatcgc acataaccga aattagaaac ttgckttcag taaacacaac 2340  
ttatacccttataatctgatta acattgcgc tcagtcgtac ccgattgtct ttaggttaact 2400  
catcgctttc caaccgtgct ctgttccgt atcagcttc gctttcggt taggttgct 2460  
gttttctgtgttatcttaca gagtagtacc taaaactact tgacttaaca ccctatctgg 2520  
gcbcacgtta tggatgtttt aggaatgtct tctcaagggc gaattc 2566

&lt;210&gt; 11

&lt;211&gt; 520

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; Beschreibung der künstlichen Sequenz: synthetic sequence coding for I-Ppo I homing-endonuclease

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (16)...(504)

&lt;400&gt; 11

gcatatgaga	tctcc	atg	gcg	ctc	acc	aat	gct	caa	atc	ttg	gct	gtg	att	51
Met		Ala	Leu	Thr	Asn	Ala	Gln	Ile	Leu	Ala	Val	Ile		

1				5				10					
---	--	--	--	---	--	--	--	----	--	--	--	--	--

gac	agt	tgg	gaa	gaa	aca	gtc	ggt	cag	ttt	cca	gtg	ata	acg	cac	cat	99
Asp	Ser	Trp	Glu	Glu	Thr	Val	Gly	Gln	Phe	Pro	Val	Ile	Thr	His	His	

15				20				25					
----	--	--	--	----	--	--	--	----	--	--	--	--	--

gta	cca	tta	ggg	ggc	ggt	ctg	caa	gga	acg	ctc	cat	tgt	tac	gag	atc	147
Val	Pro	Leu	Gly	Gly	Leu	Gln	Gly	Thr	Leu	His	Cys	Tyr	Glu	Ile		

30				35				40					
----	--	--	--	----	--	--	--	----	--	--	--	--	--

ccc	cta	gca	gct	cct	tat	ggg	gtt	ggc	ttt	gct	aag	aat	ggg	cct	acc	195
Pro	Leu	Ala	Ala	Pro	Tyr	Gly	Val	Gly	Phe	Ala	Lys	Asn	Gly	Pro	Thr	

45				50				55				60		
----	--	--	--	----	--	--	--	----	--	--	--	----	--	--

cgc	tgg	caa	tac	aaa	cg	aca	atc	aat	caa	gtc	gtc	cac	aga	tgg	ggg	243
Arg	Trp	Gln	Tyr	Lys	Arg	Thr	Ile	Asn	Gln	Val	Val	His	Arg	Trp	Gly	

65				70				75				
----	--	--	--	----	--	--	--	----	--	--	--	--

tcc	cac	aca	gtc	cct	ttt	cta	tta	gaa	ccg	gat	aac	atc	aac	ggc	aaa	291
Ser	His	Thr	Val	Pro	Phe	Leu	Leu	Glu	Pro	Asp	Asn	Ile	Asn	Gly	Lys	

80				85				90			
----	--	--	--	----	--	--	--	----	--	--	--

acc	tgc	aca	gca	tcg	cac	cta	tgt	cat	aat	act	cga	tgc	cac	aat	ccc	339
Thr	Cys	Thr	Ala	Ser	His	Leu	Cys	His	Asn	Thr	Arg	Cys	His	Asn	Pro	

95				100				105			
----	--	--	--	-----	--	--	--	-----	--	--	--

ttg	cac	ttg	tgc	tgg	gag	tca	cta	gac	gac	aac	aaa	ggc	aga	aac	tgg	387
Leu	His	Leu	Cys	Trp	Glu	Ser	Leu	Asp	Asp	Asn	Lys	Gly	Arg	Asn	Trp	

110				115				120			
-----	--	--	--	-----	--	--	--	-----	--	--	--

tgc	ccc	ggt	ccc	aac	ggg	gga	tgt	gtc	cat	gct	gtg	gtt	tgt	tta	agg	435
Cys	Pro	Gly	Pro	Asn	Gly	Gly	Cys	Val	His	Ala	Val	Val	Cys	Leu	Arg	

125				130				135				140
-----	--	--	--	-----	--	--	--	-----	--	--	--	-----

cag	ggt	ccg	ttg	tac	ggc	cca	ggg	gct	act	gtg	gca	ggt	cct	caa	caa	483
Gln	Gly	Pro	Leu	Tyr	Gly	Pro	Gly	Ala	Thr	Val	Ala	Gly	Pro	Gln	Gln	

145				150				155		
-----	--	--	--	-----	--	--	--	-----	--	--

agg ggc agt cac ttt gtg gta taactgcagc tcgagg  
 Arg Gly Ser His Phe Val Val  
 160

520

&lt;210&gt; 12

&lt;211&gt; 163

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;223&gt; Beschreibung der künstlichen Sequenz: synthetic sequence coding for I-Ppo I homing-endonuclease

&lt;400&gt; 12

Met Ala Leu Thr Asn Ala Gln Ile Leu Ala Val Ile Asp Ser Trp Glu  
 1 5 10 15Glu Thr Val Gly Gln Phe Pro Val Ile Thr His His Val Pro Leu Gly  
 20 25 30Gly Gly Leu Gln Gly Thr Leu His Cys Tyr Glu Ile Pro Leu Ala Ala  
 35 40 45Pro Tyr Gly Val Gly Phe Ala Lys Asn Gly Pro Thr Arg Trp Gln Tyr  
 50 55 60Lys Arg Thr Ile Asn Gln Val Val His Arg Trp Gly Ser His Thr Val  
 65 70 75 80Pro Phe Leu Leu Glu Pro Asp Asn Ile Asn Gly Lys Thr Cys Thr Ala  
 85 90 95Ser His Leu Cys His Asn Thr Arg Cys His Asn Pro Leu His Leu Cys  
 100 105 110Trp Glu Ser Leu Asp Asp Asn Lys Gly Arg Asn Trp Cys Pro Gly Pro  
 115 120 125Asn Gly Gly Cys Val His Ala Val Val Cys Leu Arg Gln Gly Pro Leu  
 130 135 140Tyr Gly Pro Gly Ala Thr Val Ala Gly Pro Gln Gln Arg Gly Ser His  
 145 150 155 160

Phe Val Val

&lt;210&gt; 13

&lt;211&gt; 481

&lt;212&gt; DNA

&lt;213&gt; Chlamydomonas pallidostigmatica

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (4)..(459)

&lt;223&gt; ORF coding for I-Cpa I homing endonuclease

&lt;400&gt; 13

acc atg gac att aat cct caa tgg att aca ggt ttc gta gat ggg gaa 48  
 Met Asp Ile Asn Pro Gln Trp Ile Thr Gly Phe Val Asp Gly Glu  
 1 5 10 15ggt tgt ttt agt gta agt ata ctt aga aat aat tcg ttg cgc tat ggc 96  
 Gly Cys Phe Ser Val Ser Ile Leu Arg Asn Asn Ser Leu Arg Tyr Gly  
 20 25 30cat cag ctt caa cca gaa ttc gta gtg acc caa cat aaa tta gat gca 144  
 His Gln Leu Gln Pro Glu Phe Val Val Thr Gln His Lys Leu Asp Ala  
 35 40 45

aat gtt tta tat gca tta aaa gac tac ttt aaa gtt gga tca gtc gtt 192  
 Asn Val Leu Tyr Ala Leu Lys Asp Tyr Phe Lys Val Gly Ser Val Val  
       50                 55                 60

gtg aat cat ggg gaa cgg ctt tgc tat aaa gtc aaa aat att gat cac 240  
 Val Asn His Gly Glu Arg Leu Cys Tyr Lys Val Lys Asn Ile Asp His  
       65                 70                 75

ttt ata acc gtc att ata cca ttt ttc gaa aaa cat gag cta aaa aca 288  
 Phe Ile Thr Val Ile Ile Pro Phe Phe Glu Lys His Glu Leu Lys Thr  
       80                 85                 90                 95

aaa aga aga att gaa ttt ctt cga ttt cga aaa atc tgc ttg ctg tta 336  
 Lys Arg Arg Ile Glu Phe Leu Arg Phe Arg Lys Ile Cys Leu Leu  
       100                105                110

aaa gca ggt aga cat tta gaa tcg cag gaa gga ttc gag aaa gtg ttg 384  
 Lys Ala Gly Arg His Leu Glu Ser Gln Glu Gly Phe Glu Lys Val Leu  
       115                120                125

gat tta gca aaa aaa ctc cgt atc aat gag aaa aac tac cag gaa tct 432  
 Asp Leu Ala Lys Lys Leu Arg Ile Asn Glu Lys Asn Tyr Gln Glu Ser  
       130                135                140

atc aaa cgt ttt gaa gaa act ggc gag taactcgaga gtatagagct cc 481  
 Ile Lys Arg Phe Glu Thr Gly Glu  
       145                150

<210> 14  
 <211> 152  
 <212> PRT  
 <213> Chlamydomonas pallidostigmatica

<400> 14  
 Met Asp Ile Asn Pro Gln Trp Ile Thr Gly Phe Val Asp Gly Glu Gly  
       1                 5                 10                 15

Cys Phe Ser Val Ser Ile Leu Arg Asn Asn Ser Leu Arg Tyr Gly His  
       20                25                30

Gln Leu Gln Pro Glu Phe Val Val Thr Gln His Lys Leu Asp Ala Asn  
       35                40                45

Val Leu Tyr Ala Leu Lys Asp Tyr Phe Lys Val Gly Ser Val Val Val  
       50                55                60

Asn His Gly Glu Arg Leu Cys Tyr Lys Val Lys Asn Ile Asp His Phe  
       65                70                75                80

Ile Thr Val Ile Ile Pro Phe Phe Glu Lys His Glu Leu Lys Thr Lys  
       85                90                95

Arg Arg Ile Glu Phe Leu Arg Phe Arg Lys Ile Cys Leu Leu Lys  
       100              105              110

Ala Gly Arg His Leu Glu Ser Gln Glu Gly Phe Glu Lys Val Leu Asp  
       115              120              125

Leu Ala Lys Lys Leu Arg Ile Asn Glu Lys Asn Tyr Gln Glu Ser Ile  
       130              135              140

Lys Arg Phe Glu Glu Thr Gly Glu  
       145              150

<210> 15  
 <211> 917

<212> DNA  
 <213> Chlamydomonas pallidostigmatica  
 <220>  
 <221> intron  
 <222> (18)..(893)  
 <223> CPLSU2 Intron  
 <400> 15  
 gaagacgatc ctaaggtaa ttgctagctt tagtaaaact gaactatatg ctggaaaatc 60  
 ctcgggcctc ttctttcccc agttgcacctt cgtgctaaag cgggcagatg aaactataaa 120  
 gcaaaaaagag gctcatcaact ttttttacc ggaactccgt tccggcttaa agtgatgaaa 180  
 aagatgtctt gtacgtctcg tattttgccca cgagccgtga taatcaagcg tacatgggg 240  
 caatcagcag ggaaggatgt ttgatacacg cagagtagaa ctggcagtct agaatcggtc 300  
 ggtatcgcta aacataaccc tcagagacta tacgttcgga agtctaata atcataataa 360  
 taggaggaaa ttgaaaatgg acattaatcc tcaatggatt acaggttcg tagatgggg 420  
 aggttgtttt agtctaagta tacttagaaa taattcggtt cgctatggcc atcagcttca 480  
 accagaattc gtatgtgaccc aacataaatt agatgcaa at gttttatatg cattaaaaga 540  
 ctactttaaa gttggatcaag tcggtgtgaa tcatggggaa cggcttgct ataaaagtcaa 600  
 aaatattgt cacttataaa ccgtcattat accattttc gaaaacatg agctaaaaac 660  
 aaaaagaaga attgaattc ttcgatttcg aaaaatctgc ttgctgttaa aagcaggtag 720  
 acatttagaa tcgcaggaag gattcgagaa agtgttggat ttagcaaaaa aactccgtat 780  
 caatgagaaa aactaccagg aatctatcaa acgtttgaa gaaactggcg agtaaaaaaaaa 840  
 ataagattt agatagagtc cagccttattt tgaagaaaaa tgggttaaat ctgagcgaaa 900  
 ttccttgcata agtcttc 917

<210> 16  
 <211> 21  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
 oligonucleotide primer

<400> 16  
 taaggccctc ggttagcaacg g 21

<210> 17  
 <211> 20  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
 oligonucleotide primer

<400> 17  
 ggggtaccaa atccaactag 20

<210> 18  
 <211> 25  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
 oligonucleotide primer

<400> 18  
 ggagctcgct ccccccgcgt cgttc 25

<210> 19  
 <211> 27

<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
  
<400> 19  
gatgcatgtat gacttgacgg catcctc  
  
<210> 20  
<211> 16  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
  
<400> 20  
gtcgacagat ctttaa  
  
<210> 21  
<211> 16  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
  
<400> 21  
agatctgtcg acttaa  
  
<210> 22  
<211> 22  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
  
<400> 22  
gatcccaagt taactgggt ac  
  
<210> 23  
<211> 14  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
  
<400> 23  
cccaggtaac tgga  
  
<210> 24  
<211> 25  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 24  
ttaagccagt taactggcg gagct 25 ..  
<210> 25  
<211> 17  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
<400> 25  
ccgcccagg tt aactggc 17  
<210> 26  
<211> 34  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
<400> 26  
tcgagaagat cagcctgtta tcccttagt aact 34  
<210> 27  
<211> 34  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
<400> 27  
ctagagttac tctaggata acaggctgat cttc 34  
<210> 28  
<211> 17  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
<400> 28  
agaagacgat cctaagg 17  
<210> 29  
<211> 23  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer  
<400> 29  
tgaagacttg acaaggaatt tcg 23  
<210> 30  
<211> 56  
<212> DNA  
<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 30  
agaagacat cctaaatagc aatatttacc tttgggacca aaagtttatca ggcatt 56

<210> 31

<211> 50

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 31  
tgaagacttg acaaggaatt tcgctaccc cgagtactcc aaaactaatc 50

<210> 32

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 32  
gagaagacat tcctaacaca tccataacgt gcg 33

<210> 33

<211> 32

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 33  
tgaagacttg acatttggata tggtaagta gg 32

<210> 34

<211> 175

<212> DNA

<213> Pisum sativum

<220>

<221> CDS

<222> (1)...(174)

<223> plastid translocalization sequence

<400> 34  
atg gct tct atg ata tcc tct tca gct gtg act aca gtc agc cgt gct 48  
Met Ala Ser Met Ile Ser Ser Ser Ala Val Thr Thr Val Ser Arg Ala  
1 5 10 15

tct acg gtg caa tcg gcc gcg gtg gct cca ttc ggc ggc ctc aaa tcc 96  
Ser Thr Val Gln Ser Ala Ala Val Ala Pro Phe Gly Gly Leu Lys Ser  
20 25 30

atg act gga ttc cca gtt aag aag gtc aac act gac att act tcc att 144  
Met Thr Gly Phe Pro Val Lys Lys Val Asn Thr Asp Ile Thr Ser Ile  
35 40 45

aca agc aat ggt gga aga gta aag tgc atg c 175  
Thr Ser Asn Gly Gly Arg Val Lys Cys Met

50 55

&lt;210&gt; 35

&lt;211&gt; 58

&lt;212&gt; PRT

&lt;213&gt; Pisum sativum

&lt;400&gt; 35

Met Ala Ser Met Ile Ser Ser Ser Ala Val Thr Thr Val Ser Arg Ala

1

5

10

15

Ser Thr Val Gln Ser Ala Ala Val Ala Pro Phe Gly Gly Leu Lys Ser  
20 25 30Met Thr Gly Phe Pro Val Lys Lys Val Asn Thr Asp Ile Thr Ser Ile  
35 40 45Thr Ser Asn Gly Gly Arg Val Lys Cys Met  
50 55

&lt;210&gt; 36

&lt;211&gt; 84

&lt;212&gt; PRT

&lt;213&gt; Nicotiana tabacum

&lt;220&gt;

&lt;221&gt; TRANSIT

&lt;222&gt; (1)..(84)

<223> plastidic translocalisation signal derived from  
plastidic transketolase

&lt;400&gt; 36

Met Ala Ser Ser Ser Ser Leu Thr Leu Ser Gln Ala Ile Leu Ser Arg  
1 5 10 15Ser Val Pro Arg His Gly Ser Ala Ser Ser Ser Gln Leu Ser Pro Ser  
20 25 30Ser Leu Thr Phe Ser Gly Leu Lys Ser Asn Pro Asn Ile Thr Thr Ser  
35 40 45Arg Arg Arg Thr Pro Ser Ser Ala Ala Ala Ala Val Val Arg Ser  
50 55 60Pro Ala Ile Arg Ala Ser Ala Ala Thr Glu Thr Ile Glu Lys Thr Glu  
65 70 75 80

Thr Ala Gly Ser

&lt;210&gt; 37

&lt;211&gt; 258

&lt;212&gt; DNA

&lt;213&gt; Nicotiana tabacum

&lt;220&gt;

&lt;221&gt; transit\_peptide

&lt;222&gt; (1)..(258)

<223> plastidic translocalisation signal derived from  
plastidic transketolase (frame 1)

&lt;400&gt; 37

ggtaccatgg cgtttcttc ttctctcaact ctctctcaag ctatcctctc tcgttctgtc 60

cctcgccatg gctctgcctc ttcttctcaa ctttccccctt cttctctcac tttttccggc 120

cttaaatcca atcccaatat caccacctcc cgccgcccgt a ctccttcctc cgccgccc 180

gccgcgcgtcg taaggtcacc ggcgattcgt gcctcagctg caaccgaaac catagagaaa 240

actgagactg .cgggatcc 258

<210> 38  
 <211> 260  
 <212> DNA  
 <213> Nicotiana tabacum  
 <220>  
 <221> transit\_peptide  
 <222> (1)..(260)  
 <223> plastidic translocalisation signal derived from  
       plastidic transketolase (frame 2)  
 <400> 38  
 ggtaccatgg cgtttcttc ttctctcaact ctctctcaag ctatcctctc tcgttctgtc 60  
 cctcgccatg gctctgcctc ttcttctcaa ctttcccctt cttctctcac tttttccggc 120  
 cttaaatcca atcccaatat caccacctcc cgccgcccgtatcccttcctc cgccgcccggc 180  
 gccgcccgtcg taaggtcacc ggcgattcgt gcctcagctg caaccgaaac catagagaaa 240  
 actgagactg cgctggatcc 260  
 <210> 39  
 <211> 259  
 <212> DNA  
 <213> Nicotiana tabacum  
 <220>  
 <221> transit\_peptide  
 <222> (1)..(259)  
 <223> plastidic translocalisation signal derived from  
       plastidic transketolase (frame 3)  
 <400> 39  
 ggtaccatgg cgtttcttc ttctctcaact ctctctcaag ctatcctctc tcgttctgtc 60  
 cctcgccatg gctctgcctc ttcttctcaa ctttcccctt cttctctcac tttttccggc 120  
 cttaaatcca atcccaatat caccacctcc cgccgcccgtatcccttcctc cgccgcccggc 180  
 gccgcccgtcg taaggtcacc ggcgattcgt gcctcagctg caaccgaaac catagagaaa 240  
 actgagactg cggggatcc 259  
 <210> 40  
 <211> 63  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <221> TRANSIT  
 <222> (1)..(63)  
 <223> plastidic translocalisation signal derived from  
       plastidic isopentenylpyrophosphate isomerase-2  
       (IPP-2)  
 <400> 40  
 Met Ser Ala Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser  
     1                  5                 10                 15  
 Leu Ala Leu Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro  
     20                25                 30  
 Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser  
     35                40                 45  
 Gly Thr Ala Met Thr Asp Thr Lys Asp Gly Ser Arg Val Asp Met  
     50                55                 60  
 <210> 41  
 <211> 205  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
<221> transit\_peptide  
<222> (1)..(205)  
<223> plastidic translocalisation signal derived from  
plastidic isopentenylpyrophosphate isomerase-2  
(IPP-2) (frame 1)

<400> 41  
gatatccaca ccaacaccaa tgtctgcttc ttctttattt aatctcccat tgattcgcc 60  
cagatctctc gctctttcgt ctctttttc ttctttccga tttgcccata gtcctctgtc 120  
atcgatttca ccgagaaaagt taccgaattt tcgtgcttcc tctggtaccg ctatgacaga 180  
tactaaagat ggatccccggg tcgac 205

<210> 42  
<211> 207  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> transit\_peptide  
<222> (1)..(207)  
<223> plastidic translocalisation signal derived from  
plastidic isopentenylpyrophosphate isomerase-2  
(IPP-2) (frame 2)

<400> 42  
gatatccaca ccaacaccaa tgtctgcttc ttctttattt aatctcccat tgattcgcc 60  
cagatctctc gctctttcgt ctctttttc ttctttccga tttgcccata gtcctctgtc 120  
atcgatttca ccgagaaaagt taccgaattt tcgtgcttcc tctggtaccg ctatgacaga 180  
tactaaagat ctggatcccg ggtcgac 207

<210> 43  
<211> 206  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> transit\_peptide  
<222> (1)..(206)  
<223> plastidic translocalisation signal derived from  
plastidic isopentenylpyrophosphate isomerase-2  
(IPP-2) (frame 3)

<400> 43  
gatatccaca ccaacaccaa tgtctgcttc ttctttattt aatctcccat tgattcgcc 60  
cagatctctc gctctttcgt ctctttttc ttctttccga tttgcccata gtcctctgtc 120  
atcgatttca ccgagaaaagt taccgaattt tcgtgcttcc tctggtaccg ctatgacaga 180  
tactaaagat gggatccccgg gtcgac 206

<210> 44  
<211> 44  
<212> DNA  
<213> Nicotiana tabacum

<220>  
<221> promoter  
<222> (1)..(44)  
<223> plastidic promoter PrbCL

<400> 44  
gttgcgttat atatatgaaa gagtatacaa taatgatgta tttg 44

<210> 45  
<211> 64

<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> promoter  
<222> (1)..(64)  
<223> plastidic promoter Prps16-107  
  
<400> 45  
tagcgatggg gtcttactaa agaaaaatat ttatccacct atctctatag tatatagata 60  
taga 64  
  
<210> 46  
<211> 91  
<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> promoter  
<222> (1)..(91)  
<223> plastidic promoter Prrn16  
  
<400> 46  
cgccgtcggtt caatgagaat ggataagagg ctcgtggat tgacgtgagg gggcagggat 60  
ggctatattt ctgggagcga actccggcga a 91  
  
<210> 47  
<211> 71  
<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> promoter  
<222> (1)..(71)  
<223> plastidic promoter PaccD-129  
  
<400> 47  
gtcgacatat tattttaaat aatataaagg gggttccaac atattaatat atagtgaagt 60  
gttccggatc c 71  
  
<210> 48  
<211> 27  
<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> promoter  
<222> (1)..(27)  
<223> plastidic promoter Pclpp-53  
  
<400> 48  
agacaataaa aaaaattgtt acgttac 27  
  
<210> 49  
<211> 65  
<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> promoter  
<222> (1)..(65)  
<223> plastidic promoter Prrn-62  
  
<400> 49  
gagcgaactc cgggcgaata tgaagcgcattt ggatacaagt tatgccttgg aatgaaagac 60  
aatttc 65

<210> 50  
<211> 107  
<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> promoter  
<222> (1)..(107)  
<223> plastidic promoter Prps16  
<400> 50  
tctatatata tttatataatt aaatattaat taatatgtta gataaattaa tattagcgat 60  
ggggctttac taaagaaaaa tatttatcca cctatctcta tagtata 107  
<210> 51  
<211> 47  
<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> promoter  
<222> (1)..(47)  
<223> plastidic promoter PatpB/E-290  
<400> 51  
agaaatagaa aataaagtcc aggttcgaat tccatagaat agataat 47  
<210> 52  
<211> 25  
<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> promoter  
<222> (1)..(25)  
<223> plastidic promoter PrpoB-345  
<400> 52  
aatgtgtatt atcataataa tggta 25  
<210> 53  
<211> 35  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz: consensus  
sequence of E.coli sigma70 promoter  
<400> 53  
ttgacattca ctcttcaatt atctataatg ataca 35  
<210> 54  
<211> 80  
<212> DNA  
<213> Nicotiana tabacum  
<220>  
<221> misc\_feature  
<222> (1)..(80)  
<223> 5'-untranslated region of psbA gene (incl. NcoI  
site)  
<400> 54  
tccatatttct atttgattt gtagaaaact agtgtgcttg ggagtccctg atgattaaat 60  
aaaccaagat ttaccatgg 80

<210> 55  
<211> 40  
<212> DNA  
<213> Nicotiana tabacum  
  
<220>  
<221> misc\_feature  
<222> (1)..(40)  
<223> 5'-untranslated region of rbcL gene  
  
<400> 55  
agttaggg agggattcat gacaccacaa acagccatgg 40  
  
<210> 56  
<211> 22  
<212> DNA  
<213> Nicotiana tabacum  
  
<220>  
<221> misc\_feature  
<222> (1)..(22)  
<223> 5'-untranslated region of rbcLs gene (modified)  
  
<400> 56  
agttaggg agggattcat ga 22  
  
<210> 57  
<211> 127  
<212> DNA  
<213> Synechocystis sp.  
  
<220>  
<221> misc\_feature  
<222> (1)..(127)  
<223> 3'-untranslated region of synechocystis psbA-1 gene  
  
<400> 57  
tgccattgcc ataactgctt tcggtagac ttcgtttcat ttggtaatc aaggcactc 60  
tcgcaatggg gtgccttta tggtccaagg ttaagttaa gccagtacta tttctagggt 120  
gaaatgt 127  
  
<210> 58  
<211> 90  
<212> DNA  
<213> Nicotiana tabacum  
  
<220>  
<221> misc\_feature  
<222> (1)..(90)  
<223> 3'-untranslated region of tobacco psbA gene  
  
<400> 58  
cctggctag tctataggag gtttggaaa gaaaggagca ataatcattt tcttgttcta 60  
tcaaggagggt gctattgctc ctttctttt 90  
  
<210> 59  
<211> 136  
<212> DNA  
<213> Nicotiana tabacum  
  
<220>  
<221> misc\_feature  
<222> (1)..(136)  
<223> 3'-untranslated region of tobacco rbcL gene  
  
<400> 59  
agtagacatt agcagataaa ttagcaggaa ataaagaagg ataaggagaa agaactcaag 60

taattatcct tcgttcttt aattgaattg caattaaact cggcccaatc ttttactaaa 120  
 aggattgagc cgaata 136

<210> 60  
 <211> 13  
 <212> DNA  
 <213> Künstliche Sequenz

<220>  
 <223> Beschreibung der künstlichen Sequenz: synthetic  
 ribosome binding site (RBS)

<400> 60  
 ggaggnnnnn atg 13

<210> 61  
 <211> 13  
 <212> DNA  
 <213> Künstliche Sequenz

<220>  
 <223> Beschreibung der künstlichen Sequenz: synthetic  
 ribosome binding site (RBS)

<400> 61  
 ggaggatctc atg 13

<210> 62  
 <211> 1494  
 <212> DNA  
 <213> Künstliche Sequenz

<220>  
 <223> Beschreibung der künstlichen Sequenz: complete  
 insert of vector pCB304-25

<400> 62.  
 ggtaccgggc cgataattcc tggcttagtc tataggaggt tttaaaaaga aaggagcaat 60  
 aatcattttc ttgttctatc aagagggtgc tattgtcttct ttctttttt ctcgagtacg 120  
 cttagggataa cagggtataa tagaagcttgc actctttaa agatctcgac tctatata 180  
 tttatataatt aaatattaat taatatgtta gataaattaa tattagcgat ggggtcttac 240  
 taaaagaaaaa tatttatcca cctatctta tagtatata gatctagttg tagggaggga 300  
 ttcatgacac cacaaacagc catggagct tggattgaac aagatggatt gcacgcagg 360  
 tctccggccg cttgggtgga gaggcttattc ggctatgact gggcacaaca gacaatcg 420  
 tgctctgatg ccggcgttgc cggcgttgc ggcggggc gcccgggttct ttttgtcaag 480  
 accgacctgt ccggtgcctt gaatgaactg caggacgagg cagcgcggct atcggtggctg 540  
 gccacacgg gcttcccttgc cgcagctgtc ctgcacgttgc tcaactgaagc gggaaaggac 600  
 tggctgtat tggcgaagt gcccggcag gatctctgt catctcacct tgctcctgcc 660  
 gagaaggatc ccatcatggc tgatgcaatg cggcggctgc atacgcttgc tccggctacc 720  
 tgcccattcg accaccaagc gaaacatcgc atcgagcggc cacgtactcg gatggaaaggcc 780  
 ggtcttgatc atcaggatgatc tctggacgaa gagcatcagg ggctcgcggc agccgaactg 840  
 ttgcgcaggc tcaaggcgcg catgccccgac ggcggggatc tgcgtgtac ccatggcgat 900  
 gctctgtgc cgaatatcat ggtggaaaat gcccgtttt ctggattcat cgactgtggc 960  
 cggctgggtg tggcggaccg ctatcaggac atagcgttgg ctaccctgtaa tattgctgaa 1020  
 gagctggcg gcaatgggc tgaccgttgc ctgcgtgtt acggatcgcc cgctcccgat 1080  
 tcgcagcgc tgccttcttgcctt gacgagttct tctgagcggg acccaagcta 1140  
 ggcggccgcg gatccaagct tgaattcagt agacattagc agataaattt gcaaggaaata 1200  
 aagaaggata aggagaaaga actcaagtaa ttatccttgc ttctttaat tgaattgcaa 1260  
 ttaaaactcg cccaatcttt tactaaaagg attgagccga atactcgact taaggttagcg 1320  
 tcgatcgc gcccggggag cccaaaggca cgcctggca cccgaagctc tagtatcaa 1380  
 tttggcacaat aaagcaaaat taaaatactg ataattgcca acacaattaa catctcaatc 1440  
 aaggtaaatg cttttgttttgc tttttgcacaa agctatcttgc tgcgtatcaga gctc 1494

<210> 63  
<211> 592  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz: BglII / MunI fragment from vector pCB320-192

<400> 63  
agatcttcca ttttctattt tgattttag aaaaacttagt tgcttggag tccctgatga 60  
ttaataaaac caagattttt ccatggcgct caccaatgtc caaatcttgg ctgtgattga 120  
cagttggaa gaaacagtgc gtcagttcc agtgataacg caccatgtac cattagggtgg 180  
cggtctgcaa ggaacgctcc attgttacga gatcccccta gcagctcctt atggggttgg 240  
ctttgctaag aatgggccta cccgcgtggca atacaacagg acaatcaatc aagtgcgtcca 300  
cagatgggg tcccacacag tccctttct attagaacgg gataacatca acggcaaaac 360  
ctgcacagca tcgcacactat gtcataatac tcgatgccac aatcccttgc acttgtgctg 420  
ggagtcacta gacgacaaca aaggcagaaaa ctggtgccccc ggtcccaacg ggggatgtgt 480  
ccatgcggtg gttgtttaa ggcagggtcc gttgtacggc ccaggggcga ctgtggcagg 540  
tcctcaacaa agggcagtc actttgtggt ataactgcag aagcttaat tg 592

<210> 64  
<211> 25  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 64  
aaagatctcc tcacaaaaggg ggtcg 25

<210> 65  
<211> 23  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 65  
tcgaagactt aggaccgtta tag 23

<210> 66  
<211> 26  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 66  
aggaagaccc tgcgggtaa gttccg 26

<210> 67  
<211> 31  
<212> DNA  
<213> Künstliche Sequenz  
<220>  
<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

<400> 67  
 ctcaattggg gtctctctgt ccaggtgcag g  
 <210> 68  
 <211> 672  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz: artificial fusion protein of I-Ppo I homing endonuclease and plastidic transit peptide  
 <220>  
 <221> CDS  
 <222> (1)...(669)  
 <400> 68  
 atg tct gct tct tta ttt aat ctc cca ttg att cgc ctc aga tct 48  
 Met Ser Ala Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser  
 1 5 10 15  
 ctc gct ctt tcg tct ttt tct ttc cga ttt gcc cat cgt cct 96  
 Leu Ala Leu Ser Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro  
 20 25 30  
 ctg tca tcg att tca ccg aga aag tta ccg aat ttt cgt gct ttc tct 144  
 Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser  
 35 40 45  
 ggt acc gct atg aca gat act aaa gat ggg atc ccc atg gcg ctc acc 192  
 Gly Thr Ala Met Thr Asp Thr Lys Asp Gly Ile Pro Met Ala Leu Thr  
 50 55 60  
 aat gct caa atc ttg gct gtg att gac agt tgg gaa gaa aca gtc ggt 240  
 Asn Ala Gln Ile Leu Ala Val Ile Asp Ser Trp Glu Glu Thr Val Gly  
 65 70 75 80  
 cag ttt cca gtg ata acg cac cat gta cca tta ggt ggc ggt ctgcaa 288  
 Gln Phe Pro Val Ile Thr His His Val Pro Leu Gly Gly Leu Gln  
 85 90 95  
 gga acg ctc cat tgt tac gag atc ccc cta gca gct cct tat ggg gtt 336  
 Gly Thr Leu His Cys Tyr Glu Ile Pro Leu Ala Ala Pro Tyr Gly Val  
 100 105 110  
 ggc ttt gct aag aat ggg cct acc cgc tgg caa tac aaa cgg aca atc 384  
 Gly Phe Ala Lys Asn Gly Pro Thr Arg Trp Gln Tyr Lys Arg Thr Ile  
 115 120 125  
 aat caa gtc gtc cac aga tgg ggg tcc cac aca gtc cct ttt cta tta 432  
 Asn Gln Val Val His Arg Trp Gly Ser His Thr Val Pro Phe Leu Leu  
 130 135 140  
 gaa ccg gat aac atc aac ggc aaa acc tgc aca gca tcg cac cta tgt 480  
 Glu Pro Asp Asn Ile Asn Gly Lys Thr Cys Thr Ala Ser His Leu Cys  
 145 150 155 160  
 cat aat act cga tgc cac aat ccc ttg cac ttg tgc tgg gag tca cta 528  
 His Asn Thr Arg Cys His Asn Pro Leu His Leu Cys Trp Glu Ser Leu  
 165 170 175  
 gac gac aac aaa ggc aga aac tgg tgc ccc ggt ccc aac ggg gga tgt 576  
 Asp Asp Asn Lys Gly Arg Asn Trp Cys Pro Gly Pro Asn Gly Gly Cys  
 180 185 190  
 gtc cat gcg gtg gtt tgt tta agg cag ggt ccg ttg tac ggc cca ggg 624  
 Val His Ala Val Val Cys Leu Arg Gln Gly Pro Leu Tyr Gly Pro Gly  
 195 200 205

gcg act gtg gca ggt cct caa caa agg ggc agt cac ttt gtg gta taa 672  
 Ala Thr Val Ala Gly Pro Gln Gln Arg Gly Ser His Phe Val Val  
 210 215 220

&lt;210&gt; 69

&lt;211&gt; 223

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;223&gt; Beschreibung der künstlichen Sequenz: artificial fusion protein of I-Ppo I homing endonuclease and plastidic transit peptide

&lt;400&gt; 69

Met Ser Ala Ser Ser Leu Phe Asn Leu Pro Leu Ile Arg Leu Arg Ser			
1	5	10	15

Leu Ala Leu Ser Ser Ser Phe Ser Ser Phe Arg Phe Ala His Arg Pro			
20	25	30	

Leu Ser Ser Ile Ser Pro Arg Lys Leu Pro Asn Phe Arg Ala Phe Ser			
35	40	45	

Gly Thr Ala Met Thr Asp Thr Lys Asp Gly Ile Pro Met Ala Leu Thr			
50	55	60	

Asn Ala Gln Ile Leu Ala Val Ile Asp Ser Trp Glu Glu Thr Val Gly			
65	70	75	80

Gln Phe Pro Val Ile Thr His His Val Pro Leu Gly Gly Gly Leu Gln			
85	90	95	

Gly Thr Leu His Cys Tyr Glu Ile Pro Leu Ala Ala Pro Tyr Gly Val			
100	105	110	

Gly Phe Ala Lys Asn Gly Pro Thr Arg Trp Gln Tyr Lys Arg Thr Ile			
115	120	125	

Asn Gln Val Val His Arg Trp Gly Ser His Thr Val Pro Phe Leu Leu			
130	135	140	

Glu Pro Asp Asn Ile Asn Gly Lys Thr Cys Thr Ala Ser His Leu Cys			
145	150	155	160

His Asn Thr Arg Cys His Asn Pro Leu His Leu Cys Trp Glu Ser Leu			
165	170	175	

Asp Asp Asn Lys Gly Arg Asn Trp Cys Pro Gly Pro Asn Gly Gly Cys			
180	185	190	

Val His Ala Val Val Cys Leu Arg Gln Gly Pro Leu Tyr Gly Pro Gly			
195	200	205	

Ala Thr Val Ala Gly Pro Gln Gln Arg Gly Ser His Phe Val Val			
210	215	220	

&lt;210&gt; 70

&lt;211&gt; 558

&lt;212&gt; DNA

&lt;213&gt; Physarum polycephalum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)...(555)

&lt;223&gt; coding for long version of I-PpoI

&lt;400&gt; 70

atg gcg aaa tcc aac caa gct cgg gta aac ggc ggg agt aac tat gac 48

Met Ala Lys Ser Asn Gln Ala Arg Val Asn Gly Gly Ser Asn Tyr Asp  
 1 5 10 15  
 tct ctc acc ccc tta aat atg gcg ctc acc aat gct caa atc ttg gct 96  
 Ser Leu Thr Pro Leu Asn Met Ala Leu Thr Asn Ala Gln Ile Leu Ala  
 20 25 30  
 gtg att gac agc tgg gaa gaa aca gtc ggt cag ttt cca gtg ata acg 144  
 Val Ile Asp Ser Trp Glu Glu Thr Val Gly Gln Phe Pro Val Ile Thr  
 35 40 45  
 cac cat gta cca tta ggt ggc ggt ctg caa gga acg ctc cat tgt tac 192  
 His His Val Pro Leu Gly Gly Leu Gln Gly Thr Leu His Cys Tyr  
 50 55 60  
 gag atc ccc cta gca gct cct tat ggg gtt ggc ttt gct aag aat ggg 240  
 Glu Ile Pro Leu Ala Ala Pro Tyr Gly Val Gly Phe Ala Lys Asn Gly  
 65 70 75 80  
 cct acc cgc tgg caa tac aaa cgg aca atc aat caa gtc gtc cac aga 288  
 Pro Thr Arg Trp Gln Tyr Lys Arg Thr Ile Asn Gln Val Val His Arg  
 85 90 95  
 tgg gga tcc cac aca gtc cct ttt cta tta gaa ccg gat aac atc aac 336  
 Trp Gly Ser His Thr Val Pro Phe Leu Leu Glu Pro Asp Asn Ile Asn  
 100 105 110  
 ggc aaa acc tgc aca gca tcg cac cta tgt cat aat act cga tgc cac 384  
 Gly Lys Thr Cys Thr Ala Ser His Leu Cys His Asn Thr Arg Cys His  
 115 120 125  
 aat ccc ttg cac ttg tgc tgg gag tca cta gac gac aac aaa ggc aga 432  
 Asn Pro Leu His Leu Cys Trp Glu Ser Leu Asp Asp Asn Lys Gly Arg  
 130 135 140  
 aac tgg tgc ccg ggt ccc aac ggg gga tgt gtc cat gcg gtg gtt tgt 480  
 Asn Trp Cys Pro Gly Pro Asn Gly Gly Cys Val His Ala Val Val Cys  
 145 150 155 160  
 tta agg cag ggt ccg ttg tac ggc ccg ggg gcg act gtg gca ggt cct 528  
 Leu Arg Gln Gly Pro Leu Tyr Gly Pro Gly Ala Thr Val Ala Gly Pro  
 165 170 175  
 caa caa agg ggc agt cac ttt gtg gta taa 558  
 Gln Gln Arg Gly Ser His Phe Val Val  
 180 185  
 <210> 71  
 <211> 185  
 <212> PRT  
 <213> Physarum polycephalum  
 <400> 71  
 Met Ala Lys Ser Asn Gln Ala Arg Val Asn Gly Gly Ser Asn Tyr Asp  
 1 5 10 15  
 Ser Leu Thr Pro Leu Asn Met Ala Leu Thr Asn Ala Gln Ile Leu Ala  
 20 25 30  
 Val Ile Asp Ser Trp Glu Glu Thr Val Gly Gln Phe Pro Val Ile Thr  
 35 40 45  
 His His Val Pro Leu Gly Gly Leu Gln Gly Thr Leu His Cys Tyr  
 50 55 60  
 Glu Ile Pro Leu Ala Ala Pro Tyr Gly Val Gly Phe Ala Lys Asn Gly  
 65 70 75 80

Pro Thr Arg Trp Gln Tyr Lys Arg Thr Ile Asn Gln Val Val His Arg  
 85 90 95

Trp Gly Ser His Thr Val Pro Phe Leu Leu Glu Pro Asp Asn Ile Asn  
 100 105 110

Gly Lys Thr Cys Thr Ala Ser His Leu Cys His Asn Thr Arg Cys His  
 115 120 125

Asn Pro Leu His Leu Cys Trp Glu Ser Leu Asp Asp Asn Lys Gly Arg  
 130 135 140

Asn Trp Cys Pro Gly Pro Asn Gly Gly Cys Val His Ala Val Val Cys  
 145 150 155 160

Leu Arg Gln Gly Pro Leu Tyr Gly Pro Gly Ala Thr Val Ala Gly Pro  
 165 170 175

Gln Gln Arg Gly Ser His Phe Val Val  
 180 185

&lt;210&gt; 72

&lt;211&gt; 34

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;220&gt;

&lt;221&gt; promoter

&lt;222&gt; (1) .. (34)

<223> coding for promoter derived from consensus of  
sigma70 E.coli promoters

&lt;400&gt; 72

ttgacaattt tcctctgaat tatataatta acat

34

&lt;210&gt; 73

&lt;211&gt; 413

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: artificial  
intron TetIVS2a

&lt;400&gt; 73

```

aaatagcaat attttcgtt gcctaaaaaa gttatcaggc atgcacctgg tagtagtct 60
ttaaaccaat agattgcatc ggtttaaaag gcaagaccgt caaatgcgg gaaaggggtc 120
aacagccgtt cagtaccaag tctcaggaga aactttgaga tggccttgca aagggtatgg 180
taataagctg acggacatgg tcctaaccac gcagccaagt cctaagtcaa cagatcttct 240
gtttagatgg atgcagttca cagactaaat gtccgtcggg gaagatgtat tttctcata 300
agatatagtc ggacctctcc ttaatggag cttagcgatg aagtgtatgca acactggagc 360
cgctggAAC taatttgtat gcgaaagtat attgattgtat tttggagtac tcg 413

```

&lt;210&gt; 74

&lt;211&gt; 791

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: insert of  
plasmid pCB459-1

&lt;400&gt; 74

ttacaatttc cattcgccat tcaggctgctgcaactgttgg gaagggcgat cgggtgcgggc 60

ctcttcgctta	ttacggccagc	tggcgaaagg	gggatgtgtc	gcaaggcgat	taagtgggt	120
aacgccagg	tttccccagt	cacgacgttg	taaaaacgacg	gccagtgagc	gcgcgtaata	180
cgactcacta	tagggcgaat	tgggtaccgg	gcccccccctc	gagggtcgacg	gtatcgataa	240
gcttgacccg	acaaggaaatt	tcgctcgagt	actccaaaac	taatcaatat	actttcgcat	300
acaaaattagt	tcccagcggc	tccagtgttg	catcaactca	tccgctagct	cccattaagg	360
agaggtccga	ctatatctta	tgagaagaat	acatcttccc	cgaccgacat	ttagtctgtg	420
aactgcatcc	atatcaacag	aagatctgtt	gacttaggac	ttggctgcgt	ggttaggacc	480
atgtccgtca	gcttattacc	ataccctttg	caaggccatc	tcaaagttc	ccctgagact	540
tggtaactgaa	cggctgttga	cccccttccc	gcaatttgac	ggtcttgctt	tttaaaccga	600
tgcaatctat	tggtttaaag	actagctacc	aggtgcatgc	ctgataactt	tttaaggcaa	660
cgaaaaaatat	tgcttattac	cttaggaccc	ttaggatcca	ctagttctag	agcggccgcc	720
accgcgggtgg	agctccagct	tttgttccct	ttagtgaggg	ttaattgcgc	gcttggcgta	780
atcatggtca	t					791

<210> 75

<211> 1291

<212> DNA

## <213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insert of vector pCB478-3

<400> 75

ttaacaatttc	cattcgccat	tcaggctgct	caactgttgg	gaagggcgat	cggtgcggc	60
ctcttcgcta	ttacgcccagc	tggcgaaagg	gggatgtgct	gcaaggcgat	taagtgggt	120
aacgccaggg	tttcccagt	cacgacgtt	taaaaacgacg	gccagtgagc	gcgcgtaata	180
cgactctacta	tagggcgaat	tgggtaccgg	gccccccctc	gagggtcgacg	gtatcgataa	240
gcttgaccgg	acaaggaatt	tcgctcgagt	actccaaaac	taatcaatat	actttcgcat	300
acaaaattagt	tcccagcgcc	tccagtgtt	catcaactca	tccgctagct	cccattaagg	360
agaggtccga	ctatatctta	tgagaagaat	tttgatcctt	actcgccagt	ttcttcaaaa	420
cgttttagat	gttccggta	gtttttctca	ttgtatacgga	gttttttgc	taaatccaac	480
actttctcga	atccttcctg	cgattctaaa	tgtctacctg	cttttaacag	caagcagatt	540
tttgc当地	gaagaaattc	aattcttctt	tttgc当地	gctcatgtt	ttcgaaaaat	600
ggtataatga	cgggtataaaa	gtgatcaata	tttttgactt	tatagcaaag	ccgttccccca	660
tgattcacaa	cgactgatcc	aactttaaag	tagtcttttta	atgcataaa	aacatttgca	720
tctaatttat	gttgggtcac	tacgaattaa	ttctgggttga	agctgatggc	catagcgcaa	780
cgaatttattt	ctaagtatac	ttacactaaa	acaacaccttcc	ccatctacga	aacctgtataat	840
ccatttgagga	ttaatgtcca	ttgtccctcc	ctacaacgtc	tgcagatcaa	acatcttccc	900
cgaccgacat	ttagtctgt	aactgcacatcc	atatcaaacag	aagatctgtt	gacttaggac	960
ttggctgcgt	ggttaggacc	atgtccgtca	gcttattacc	atacccttgc	caaggccatc	1020
tcaaaggttt	ccctgagact	tggtaactgaa	cggctgttga	cccccttccc	gcaatttgac	1080
ggtcttgcct	tttaaaccga	tgcaatctat	tggttaaag	actagctacc	aggtgcatgc	1140
ctgataactt	tttaaggcaa	cgaaaaatat	tgcttattac	cttaggaccg	tttagatccca	1200
ctagttctag	agccggccgccc	accgcggtgg	agctccagct	tttgc当地	tttagtgagg	1260
ttaatttgccg	gcttggcgta	atcatggtca	t			1291

<210> 76

<211> 1386

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: insert of vector pCB492-25

<400> 76

```

gcgcgctaa tacgactcac tataggcga attggggcg actgttacc aaaaacacag 60
gtctccgcaa agtcgtaaga ccatgtatgg gggctgacgc ctgcccagtg cccgaaggtc 120
aaggaaagtgtg gtgacacctgat gacaggggag cccgcgaccg aagccccgtt gaacggcggc 180
cgtaactata acggtcctaa ggttaaatagc aatatttttc gttgccttaa aaagtttatca 240

```

ggcatgcacc tggtagctag tctttaacc aatagattgc atcggttaa aaggcaagac 300  
 cgtcaaattg cgggaaaggg gtcaacagcc gttcagtacc aagtctcagg gaaaactttg 360  
 agatggcctt gcaaagggtt tggtataag ctgacggaca tggtcctaac cacgcagcca 420  
 agtcctaagt caacagatct tctgtgata tggatgcagt tcacagacta aatgtcggtc 480  
 ggggaagatg tttgatctgc agacgttga gggagggacc atggacatta atcctcaatg 540  
 gattacaggt ttcgtatgtt gggaaagggtt ttttagtgta agtataactta gaaataattc 600  
 gttgcgtat ggccatcagc ttcaaccaga attcgtatgt acccaacata aatttagatgc 660  
 aaatgttta tatgcattaa aagactactt taaagttgga tcagtcgtt tgaatcatgg 720  
 ggaacggctt tgctataaaag tcaaaaaat tgcatactt ataaccgtca ttataccatt 780  
 tttcgaaaaa catgagctaa aaacaaaaag aagaattgaa ttcttcgtat ttcgaaaaat 840  
 ctgcttgctt taaaagcgag gtagacattt agaatcgag gaaggattcg agaaagtgtt 900  
 ggatttagca aaaaaactcc gtatcaatga gaaaaactac caggaatcta tcaaacgttt 960  
 tgaagaaaact ggcgagtaag gatcaaaattt cttctctataa gatatagtcg gacctctcct 1020  
 taatgggagc tagcggatga agtgcataa cactggagcc gctggaaact aatttgcatt 1080  
 cggaaagtata ttgatttagtt ttggagtact cgagcgaat ttcttgcgg gtaagttccg 1140  
 acccgcacga aaggcgtaac gatctggca ctgtctcgga gagaggtcg gtgaaataga 1200  
 catgtctgtt aagatgcggc ctaccgtcac ctggacagag agaccccaat ttctggccta 1260  
 gtctatagga ggttttggaaa agaaaggagc aataatcatt ttcttgcgtt atcaagaggg 1320  
 tgctattgct cctttctttt ttctcgacc agctttgtt ccctttagtg agggtaattt 1380  
 gcgcgc 1386

&lt;210&gt; 77

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

&lt;400&gt; 77

tagtaaatga caatttcct ctgaattata taattaacat ggcgactgtt tacaaaaaac 60

&lt;210&gt; 78

&lt;211&gt; 31

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

&lt;400&gt; 78

ctcaatttggg gtctctctgt ccaggtgcag g

31

&lt;210&gt; 79

&lt;211&gt; 1258

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: PCR product  
Prom-TetIVS2a-Cpa

&lt;400&gt; 79

 tagtaaatga caatttcct ctgaattata taattaacat ggcgactgtt tacaaaaaac 60  
 acaggctcc gcaaagtctgt aagaccatgt atggggctg acgcctgccc agtgcggaa 120  
 ggtcaaggaa gttgggtgacc tgatgacagg ggagccggcg accgaagccc cggtaacgg 180  
 cggccgttaac tataacggc ctaaggtaaa tagcaatatt ttctcggtt taaaaagggtt 240  
 atcaggcatg cacctggtag ctatcttta aaccaataga ttgcgttgcgt taaaaggca 300  
 agacgtcaa attgcgggaa aggggtcaac agccgttcaag taccatgtt cagggaaac 360  
 tttagatgg ccttgcaaaag ggtatggtaa taagctgacg gacatggtcc taaccacgca 420  
 gccaagtcc aagtcaacag atcttcgtt gatatggatg cagttcacag actaaatgtc 480

ggtcggggaa gatgtttgat ctgcagacgt ttagggagg gaccatggac attaatcctc 540  
 aatggattac aggttgcga gatggggaa gttgttttag tgtaagtata cttagaaata 600  
 attcgttgcg ctatggccat cagcttcaac cagaattcgat agtacccaa cataaattag 660  
 atgcaaatgt ttatatgcata taaaagact actttaaagt tggatcgtc gttgtgaatc 720  
 atggggAACG GCTTGCAT aaagtaaaa atattgatca ctttataacc gtcattatac 780  
 cattttcga aaaacatgag ctaaaaacaa aaagaagaat tgaatttctt cgatttcgaa 840  
 aaatctgctt gctgttaaaa gcaggttagac atttagaattc gcaggaagga ttcgagaaag 900  
 tggatcgtt agcaaaaaaa ctccgtatca atgagaaaaaa ctaccaggaa tctatcaaac 960  
 gtttgaaga aactggcgag taaggatcaa aattcttctc ataagatata gtcggaccc 1020  
 tccttaatgg gagctagcgatgatgat gcaacactgg agccgctggg aactaatttg 1080  
 tatgcgaaag tatattgatt agtttggag tactcgagcg aaattcttg tcgggtaagt 1140  
 tccgacccgc acgaaaggcg taacgatctg ggcactgtct cggagagagg ctccgtgaaa 1200  
 tagacatgtc tgtgaagatg cggactaccc gcacctggac agagagaccc caatttag 1258

&lt;210&gt; 80

&lt;211&gt; 4746

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; Beschreibung der künstlichen Sequenz: insert of vector pCB435-45

&lt;400&gt; 80

gcgcacaata cgcaaacgc ctctccccgc gctgtggccg attcattaat gcagtatgat 60  
 cttgacggca tcctcacctt cctccggctt atcaccggca gtctgttcag ggttccaaac 120  
 tcaacatgg caactaaaca cgagggttgc gtcgttgcg ggacttaacc caacaccc 180  
 cggcacgacgc tgacgacacgc catgcaccac ctgtgtccgc gttccgaaag gcacccctct 240  
 ctttcaagag gattcgcggc atgtcaagcc ctggtaaggt tcttcgtttt gcattcgaaatt 300  
 aaaccacatg ctccaccgcgt tgcggggcccccgtcaatt ctttttagtt tcatttttgc 360  
 gaacgtactc cccaggcggg atacttaacg cgtagctac agcactgcac gggtcgatac 420  
 gcacagcgcc tagtatccat cgtttacggc taggactact ggggtatcta atcccattcg 480  
 ctccccctagc ttctgtctct cagtgtcagt gtcggcccgag cagagtgcctt tcggcgttgg 540  
 tttttttcc gatctctacg catttaccgc ctccaccggaa aattccctct gccccttaccg 600  
 tactccagct tggtagtttcc caccgcgtt ccagggttgc gcccggat ttgacggcg 660  
 actaaaaag ccacctacag acgtttacgc cccaaatcatt cggataacgc ttgcatact 720  
 ctgtattacc gggctgtcg gcacagagtt agccgatgc tattccccag ataccgtcat 780  
 tgcctttctt cccggaaaag aagtccacgc cccgtggcc ttctacctcc acgcggcatt 840  
 gtcggcgtcag gtttcgcccc attgcccggaa attccccact gtcctccc gtggggatc 900  
 gggccgtgtc tcaactcccg tgggtgtat catttcgtc gaccagctac tgatcatcg 960  
 cttggtaagc tattgcctca ccaactagct aatcagacgc gagccctcc tcggggcgat 1020  
 tcctccctttt gtcctcgtc ctacgggttgc ttgcggccg tttccagctg ttgtttccct 1080  
 cccaaggggca ggttcttacag cggtactacac ccgtccggca ctggaaacac cactttccgt 1140  
 ccgacttgc tggtaagc atgcccggc cgttcatctt gagccaggat cgaactctcc 1200  
 atgagattca tagttgcatt acttatacg tccctgttgc tagacaaagc ggattcggaa 1260  
 ttgttttca ttccaaggca taacttgtat ccatgcgtt catattgcgc cggagttcg 1320  
 tcccagaaat atagccatcc ctggcccttc acgtcaatcc cacgagcctc ttatccattc 1380  
 tcattgaacg acggcgcccccccccgtcgc atatattta aataatataa 1440  
 agggggttcc aacatattaa tatatacgat ggttccgg atccactagg acgtttagg 1500  
 gagggaccat ggacattaat cctcaatggat ttacagggtt cgttagatggg gaaggttgg 1560  
 ttgtgtaaatataacttgcgt tgccgtatgg ccatcagctt caaccagaat 1620  
 tcgttagtgc ccaacataaa ttagatgcaat gttttataa tgcataaaaa gactacttta 1680  
 aagttggatc agtcgttgcg aatcatgggg aacggctttg ctataaagtc aaaaatattg 1740  
 atcactttat aaccgtcattt ataccatttt tcgaaaaacaa tgacgtaaaaa aaaaaaagaa 1800  
 gaattgaatt tttcgattt cggaaaaatctt gttgtctgtt aaaagcaggat agacatttag 1860  
 aatcgcggaa aggattcgag aaagtggg atttagcaaa aaaaactccgt atcaatgaga 1920  
 aaaaactacca ggaatctatc aaacgttttgg aagaaaactgg cgagtaactc gagaaaaaaa 1980  
 gaaaggagca atagccaccc tttgtatgaa caagaaaaatg attattgtctt ctttctttc 2040  
 aaaaacccctt atagactagg ccaggatattc tctgcagttt tttgccaactt accttagtga 2100

tctcgccccc cacgtagtgg acaaattctt ccaactgatc tgccgcgcag gccaaacgat 2160  
 cttcttcgg tccaagataa gcctgtctag cttcaagtat gacgggctga tactgggccc 2220  
 gcaggcgctc cattgcccaag tcggcagcga cttcccttcgg cgccgattttgc ccgggttactg 2280  
 cgctgtacca aatgcgggac aacgtaaagca ctacatttcg ctcatcgcca gcccacttcgg 2340  
 gccccggagtt ccatagcggt aagggttcat tttagcgccctc aaatagatcc ttgtcaggaa 2400  
 ccggatcaaa gagttccctcc gcccgtggac ctaccaaggc aacgtatgt tctcttgctt 2460  
 ttgtcagcaa gatagccaga tcaatgtcg tctgtggctgg ctcgaagata cctgcaagaa 2520  
 tgtcatttgcg ctggcatttc ccaaatttgcg gttcgccgtt agctggataa cggccacggaa 2580  
 tgatgtcgctc gtgcacaaca atgggtactt ctacagcgcg gagaatctcg ctctctccag 2640  
 gggaaagccga agtttccaaa aggtcggttga tcaaaagctcg ccgcgttgtt tcatcaagcc 2700  
 ttacggtcac cgtaaccaggc aaatcaatat cactgtgtgg cttcaggccg ccatccactg 2760  
 cggagccgtc caaatgtacg gccagcaacg tcgggttcggat atggcgctcg atgacgcca 2820  
 ctacctctga tagttgagtt gatacttcgg cgataaccgc ttacagagcc atgaatccct 2880  
 ccctacaact agatcctcgcc ccggagttcg ctcccagaaa tatagccatc cctggccccc 2940  
 cacgtcaatc ccacgagccctt cttatccatt ctcatgttgcg aacggcgctcg aaaaaaaaaa 3000  
 cggtaaaaaa tccaaactaga aaaactcaca ttgggcttag ggataatcag gctcgaaactg 3060  
 atgacttcca ccacgtcaag gtgacactt accgctgagt tatatccctt ccccgccccca 3120  
 tcgagaaaata gaactgacta atcctaagtc aaagggtcgaa gaaactcaac gccactattc 3180  
 ttgaacaact tggagccggg ctttcttttc gcactattac ggatatgaaa ataatggtca 3240  
 aaatccggatt caattgtcaa ctggccctat cggaaatagg attgactacc gattccgaag 3300  
 gaactggagt tacatctttt ttccattcaa gagttcttat gcgttccac gcccctttga 3360  
 gaccccgaaa aatggacaaa ttccctttct taggaacaca tacaagattc gtcactacaa 3420  
 aaaggataat ggtaacccta ccattaacta cttcatttat gaatttcata gtaatagaaa 3480  
 tacatgtcct accgagacag aatttggaaat ttgctatcctt cttgccttagc aggccaaagat 3540  
 ttaccccggtt ggaaaggatg attcatttcgg atcgacatga gagtccaaact acattggccag 3600  
 aatccatgtt gtatatttga aagaggttga cttccttgct tctctcatgg tacactcctc 3660  
 ttcccgccga gccccttttc tcctcggtcc acagagacaa aatgttaggac tggtgccaa 3720  
 aattccatcg actcactaag tcgggatcac taactaata taatctaata taatagtcta 3780  
 atatatctaa tataatagaa aatactaata taatagaaaaa gaactgtctt ttctgtatac 3840  
 tttcccggtt tccgttgcta ccgaggccct cgtgatacgc ctattttat aggttaatgt 3900  
 catgacccaaa atcccttaac gtgagtttc gttccactga gctgcagacc ccgtagaaaaa 3960  
 gatcaaaggaa tcttctttagt atcctttttt tctgcgcgtat atctgctgtc tgcaaaacaaa 4020  
 aaaaccacccg ctaccagcg gggtttgttt gcccgttcaaa gagctacccaa ctctttttcc 4080  
 gaaggttaact ggcttcagca gagcgcagat accaaataact gtcctcttag tctgtccgt 4140  
 gtttaggcccac cacttcaaga actctgttagc accgcctaca tacctcgctc tgctaatcct 4200  
 gttaccaggcgtt gctgtccatc gtggcgataa gtcgtgtctt accgggttgg actcaagacg 4260  
 atagttaccg gataaggcgcc agcggctggg ctgaacgggg gtttcgtgcac cacagcccg 4320  
 cttggagcgca acgacactaca ccgaactgag atacctacag cgtgagctat gagaagcg 4380  
 cacgcttccc gaagggagaa aggccgacag gtatccggta agccgcaggg tcggAACAGGG 4440  
 agagcgacg agggagcttc cagggggaaa cggctggat ctttatagtc ctgtcggtt 4500  
 tcgccccctc tgacttgcgtt gtcgatccatc gtatgtctcg tcaggggggc ggagccatcg 4560  
 gaaaaacgcc agcaacgcgg ctttttacg gttccctggcc ttttgcgttcc cttttgctca 4620  
 catgttctttt cctgcgttcatc cccctgattt ttttgcgttcc cttttgatgtt 4680  
 agctgataacc gtcgcccgc gccgaacgac cgagcgac gactgtga gcgaggaagc 4740  
 ggaaga 4746

&lt;210&gt; 81

&lt;211&gt; 1213

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: probe for  
Southern blot analysis (16SrDNA)

&lt;400&gt; 81

ttatcaccgg cagtctgttc aggggtccaa actcaacgat ggcaactaaa cacgagggtt 60  
 gcgctcggtt cgggacttaa cccaaacacct tacggcgcga gctgacgaca gccatgcacc 120  
 acctgtgtcc gctgtcccgaa aggcccccct cttttcaag aggattcgcg gcatgtcaag 180

ccctggtaag gttcttcgt ttgcattcgaa ttaaaccaca tgctccaccc cttgtgcggg 240  
 cccccgtcaa ttccttttag tttcattctt gcgAACgtac tccccaggcg ggataactaa 300  
 cgcgtagct acagcaactgc acgggtcgat acgcacagcg cctagtatcc atcgtttaacg 360  
 gctaggacta ctggggtatac taatcccatt cgctccccta gctttcgct ctcagtgtca 420  
 gtgtcgcccc agcagagtgc ttccgggtt ggtttctt ccgatctcta cgcatttcac 480  
 cgctccaccc gaaattccct ctgcccctac cgtactccag cttggtagtt tccaccgcct 540  
 gtccagggtt gagccctggg atttgacggc ggacttaaaa agccacctac agacgctta 600  
 cgcccaatca ttccggataa cgcttgcata ctctgtatta ccgcggctgc tggcacagag 660  
 ttagccatg ctatccccc agataccgtc attgcttctt ctccggaaa agaagttcac 720  
 gaccgtggg ccttctaccc ccacgcggca ttgctccgtc aggcttcgc ccattgcggga 780  
 aaattccca ctgctgcctc ccgttaggat ctggggcgtg tctcagtccc agtgtggctg 840  
 atcatccctc cggaccagct actgatcata gccttggtaa gctattgcct caccactag 900  
 ctaatcagac gcgagcccc cctcggccgg attcctcctt ttgctctca gcctacgggg 960  
 tattagcagc cgtttccagc tttttttccc ctcccaaggg cagggttctt cgcgttactc 1020  
 acccgccgc cactggaaac accacttccc gtccgacttg catgtgttaa gcatgccccc 1080  
 agcggtcatac, ctgagccagg atcgaactct ccatgagatt catagttgca ttacttata 1140  
 ctcccttgtt cgtagacaaa gcggattcgg aattgtctt cattccaagg cataacttgt 1200  
 atccatgcgc ttc 1213

&lt;210&gt; 82

&lt;211&gt; 949

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: probe for  
Southern blot (23SrDNA)

&lt;400&gt; 82

catcgctcgag agggaaacag cccggatcac cagctaaggc ccctaaatga tcgctcagtg 60  
 ataaaaggagg taggggtgca gagacagcca ggagggttgc ctagaagcag ccacccttga 120  
 aagagtgcgt aatagctcac tgatcgacgct ctcttgcgcc gaagatgaac ggggctaagc 180  
 gatctggcga agctgtggg tggtaaaaata catcggtagg ggagcggttcc gccttagaga 240  
 gaagcctccg cgcgagcggt ggtggacgaa gcggaaagcga gaatgtcgcc ttgagtaacg 300  
 caaacattgg tgagaatcca atgccccgaa aacctaaggg ttccctccgca aggttcgtcc 360  
 acggagggtg agtcaggccc taagatcagg ccgaaaggcg tagtgcgttgg acaacaggtg 420  
 aatattccctg tactgcccct tttttgtccc gagggacgga ggaggctagg tttagccgaaa 480  
 gatggttatc ggttcaagaa cgtaaagggtt ccctgttttgc ttagggtaag aaggggtttaga 540  
 gaaaatgcct cgagccaatg ttcaataacc aggcgctacg gcgctgaagt aaccatgcc 600  
 atactccctc gaaaagctcg aacgactttg agcaagaggg tacctgttacc cgaaaccgcac 660  
 acaggtgggt aggttagagaa tacctagggg cgcgagacaa ctctctctaa ggaactcgcc 720  
 aaaatagccc cgtaaacttcg ggagaagggg tgcctccctca caaagggggt cgcagtgacc 780  
 aggcccccggc gactgtttac caaaaacaca ggtctccgca aagtgcgttgg accatgtatg 840  
 ggggctgacg cctgcccagt gccggaaagggtt caaggaagttt ggtgacactga tgacagggga 900  
 gcccggcacc gaagccccgg tgaacggccgg ccgttaactat aacggtcct 949

&lt;210&gt; 83

&lt;211&gt; 1287

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: insert of  
vector pCB456-2

&lt;400&gt; 83

gagctctcag tactcgagac atttcaccct agaaatagac ttaaacttta ctggcttaac 60  
 ttttaaccttg gaccataaaa ggcaccccat tgcgagagtg cccttgattt accaaatgaa 120  
 acgaagtcta accgaaagca gttatggcaa tggcagaatt ctgatcacgg aagatagtt 180  
 tggcaaaaaa agcaaaaagc atttaccttg attgagatgt taattgtgtt ggcaattatc 240  
 agtatattaa ttttgctttt tttgtccaaat ttgataactag agcttcgggt gccagggcg 300

gcccgggc tccccgggcg cgtactcgac gctaccaa gagagtcaag ctaattctaa 360  
 ttcgcgtct agcgtgcggc cgctctagaa ctatggatc ccccggtcg cagttattt 420...  
 ccaactacct tagtgatctc gccttcacg tagtgacaa attctccaa ctgatctcg 480  
 cgcgaggcca agcgatctc ttcttgtcca agataagcct gtcttagttc aagtatgacg 540  
 ggctgatact gggccggcag ggcgtccatt gcccagtcgg cagcgacatc cttcgcccg 600  
 attttccgg tactgcgt gtaccaa atg cgggacaacg taagcactac atttcgctca 660  
 tcgcagccc agtcggccgg cgagttccat agcgtaagg tttcatttag cgcctcaaat 720  
 agatctgtt caggaaccgg atcaaagagt ccctccgccc ctggacctac caaggcaacg 780  
 ctatgttctc ttgcgtttgt cagcaagata gccagatcaa tgtcgatcg gtgcggctcg 840  
 aagataacctg caagaatgtc attgcgtgc cattctccaa attgcgttc ggcgttagct 900  
 ggataacgcc acggaatgt gtcgtgtgc acaacaatgg tgacttctac agcgcggaga 960  
 atctcgctct ctccagggga agccgaagtt tccaaaaggt cgttgatcaa agctcgccgc 1020  
 gttttcat caaggcttac ggtcacccgt accagcaaat caatatcaat gtgtggcttc 1080  
 aggcccat ccactgcggc gccgtacaaa tgtacggcca gcaacgtcg ttcgagatgg 1140  
 cgctcgatga cgccaaactac ctctgatagt tgagttgata cttcgccat aaccgcttca 1200  
 cgagccatga gatcctccag atccatgtat cattatagat aattgaagag tgaatgtcaa 1260  
 gtcgacctcg agggggggcc cggtacc 1287

&lt;210&gt; 84

&lt;211&gt; 1387

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; Beschreibung der künstlichen Sequenz: insert of vector pCB528-2

&lt;400&gt; 84

ggtaccccg ttaactggag atctcgagga ggtcacatgg gagctggat tgaacaagat 60  
 ggattgcacg cagggtctcc ggccgttgg gtggagaggc tattcggcta tgactggca 120  
 caacagacaa tcggctgctc tgatggccgc gtgttccggc tgtcagcgca ggggccccg 180  
 gttcttttg tcaagaccga cctgtccggt gcccgtaaatg aactgcagga cgaggcagcg 240  
 cggctatcggt ggctggccac gacggcggtt cttgcgcag ctgtgtcgta cttgtcaact 300  
 gaagcggaa gggactggct gctattggc gaagtgcggc ggcaggatct cctgtcatct 360  
 cacctgctc ctgcccggaa agtattccatc atggctgatg caatgcggcg gtcgcatacg 420  
 cttgatccgg ctacctgccc attgcaccac caagcgaaac atgcgtcgta gcgagcacgt 480  
 actcgatgg aagccggctc tgtcgatcg gatgatctgg acgaagagca tcaggggctc 540  
 ggcgcggcc aactgttcgc caggctcaag ggcgcgcgtc ccgacggcgatgc 600  
 gtgaccatg gcgatgcctg cttgccaat atcatggatgg aaaatggccg cttttctgga 660  
 ttcatcgact gtggccggct ggggtggcg gaccgtatc aggacatagc gttggctacc 720  
 cgtgatattg ctgaagagct tggcgccaa tgggctgacc gcttcctcgat gctttacgg 780  
 atcgcgcgtc cgcattcgca ggcgcgtcc ttctatcgcc ttcttgacga gttttctgta 840  
 tctagatagg aggtcatcat ggcgcgtacc aatgctcaaaa tcttggctgt gattgacagt 900  
 tgggaagaaa cagtcggcgtca gtttcaggta ataacgcacc atgtaccatt aggtggcggt 960  
 ctgcaaggaa cgctccattt ttacgagatc cccctagcg cttcttatgg gttggcttt 1020  
 gctaagaatg ggcctacccg ctggcaatac aaacggacaa tcaatcaagt cgtccacaga 1080  
 tgggggtccc acacagttcc ttttcttata gaaccggata acatcaacgg caaaacctgc 1140  
 acagcatcgcc acctatgtca taatactcgta tgccacaatc cttgcactt gtgctggag 1200  
 tcactagacg acaacaaagg cagaaactgg tgccccggtc ccaacggggg atgtgtccat 1260  
 gcggtggttt gtttaaggca gggtcgttg tacggcccgag gggcgactgt ggcaggtcct 1320  
 caacaaaggg gcaagtcactt tttttttttt tttttttttt tttttttttt tttttttttt 1380  
 ggagctc 1387

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**